1969

Environmental correlations between paternal half-sibs for milk and milk fat production

Krishan Kumar Arora
Iowa State University

ENVIRONMENTAL CORRELATIONS BETWEEN PATERNAL HALF-SIBS FOR MILK AND MILK FAT PRODUCTION

by

Krishan Kumar Arora

A Dissertation Submitted to the
Graduate Faculty in Partial Fulfillment of
The Requirements for the Degree of
DOCTOR OF PHILOSOPHY

Major Subject: Animal Breeding

Approved:

Signature was redacted for privacy.

In Charge of Major Work

Signature was redacted for privacy.

Head of Major Department

Signature was redacted for privacy.

Dean of Graduate College

Iowa State University
Of Science and Technology
Ames, Iowa

1969

TABLE OF CONTENTS

		Page
I.	INTRODUCTION	1
II.	REVIEW OF LITERATURE	5
III.	SOURCE AND ADJUSTMENT OF DATA	19
IV.	DERIVATION OF REGRESSIONS	22
	A. Regression of the Average of Daughters in a Sample on the Average of Daughters in Another Sample	23
	B. Regression of Breeding Value on the Daughter Average	25
V.	RESULTS AND DISCUSSION	29
	A. Heritability	29
	B. Environmental Correlations	34
	1. The difference in intra-class correlations among daughters in one herd versus in many herds	34
	2. Intra-class correlation due to herds, year-seasons and herd-year-seasons	40
	3. Within herd and sire intra-class correlation	51
	4. Simultaneous estimation of c^2 and h^2	56
VI.	SUMMARY	77
VII.	LITERATURE CITED	80
VIII.	ACKNOWLEDGMENTS	83

LIST OF FIGURES

		Page
Figure 1.	The repeatability of the progeny test based on varying number of daughters	7
Figure 2.	A theoretical example of regressions	16
Figure 3.	Path diagram illustrating the correlation of breeding value on the daughter average	27
Figure 4.	Convergence of least squares estimates of c^2 and h^2 for deviation milk in various cycles of iteration (before adjustment)	67
Figure 5.	Convergence of least squares estimates of c^2 and h^2 for deviation milk fat in various cycles of iteration (before adjustment)	68
Figure 6.	Convergence of least squares estimates of c^2 and h^2 for deviation milk in various cycles of iteration (after adjustment)	69
Figure 7.	Convergence of least squares estimates of c^2 and h^2 for deviation milk fat in various cycles of iteration (after adjustment)	70
Figure 8.	Convergence of least squares estimate of h^2 for deviation milk in various cycles of iteration (before adjustment)	71
Figure 9.	Convergence of least squares estimate of h^2 for deviation milk fat in various cycles of iteration (before adjustment)	72
Figure 10.	Convergence of least squares estimate of h^2 for deviation milk in various cycles of iteration (after adjustment)	73
Figure 11.	Convergence of least squares estimate of h^2 for deviation milk fat in various cycles of iteration (after adjustment)	74

LIST OF TABLES

		Page
Table 1.	Expected values for $r_{G_S(D\text{-}HA)}$ with $h^2 = .25$, $s = .32$, $k = 10$, and $c^2 = .00$, .01, .02, .04, and .06	11
Table 2.	Analysis of variance - heritability estimates	32
Table 3.	Estimates of heritability	33
Table 4.	Analysis of variance - heritability and correlation among half-sibs in the same herd	36
Table 5.	Analysis of variance - heritability and correlation among half-sibs in different herds	37
Table 6.	Analysis of variance - single herd sires only	38
Table 7.	Estimation of environmental correlation among half-sibs from heritability estimates obtained under situations a) and b)	39
Table 8.	Estimates of intra-class correlations obtained under situations a) and b)	40
Table 9.	Components of covariance among paternal half-sibs in the same and different herds and year-seasons	42
Table 10.	Analysis of variance - correlation (s) among cows in the same herd and year-season	45
Table 11.	Analysis of variance - correlation (t) among cows in the same herd but different year-seasons	46
Table 12.	Analysis of variance - correlation (w) among cows in the same year-season but in different herds	47
Table 13.	The additional correlations (ρ^2) among paternal half-sisters in the same herd	48
Table 14.	Analyses of variance - intra-herd correlation among daughters of AI sires and NS sires (deviation records)	54
Table 15.	Estimates of variance components and the intra-herd correlation (deviation records)	55
Table 16.	Values of variables involved in equation for regression of breeding value on daughter average	63

Page

Table 17. Least squares estimates of h^2 and c^2 obtained by using Hartley's modified Gauss-Newton method 64

I. INTRODUCTION

The importance of artificial insemination (AI) in dairy cattle breeding has increased considerably in recent years and is now one of the major tools for producing genetic improvement. Since the beginning of artificial insemination, the number of daughters in a progeny test and the distribution of daughters over herds have been a serious problem in AI sire proofs. In the early history of AI, studs had to depend on naturally proven bulls with single herd proofs for selection of their sires. When a substantial number of naturally proven bulls were used in AI service and their daughters production became available, the repeatability of natural service proofs in AI was found to be rather low. Then the problem of the distribution of a bull's daughters over herds became acute and meaningful.

Several possible causes for the low correlations between the estimates of breeding values based on single-herd proofs (natural proofs) and those based on extensive multiple proofs (AI proofs) can be suggested, viz.:

1. Considering the lack of repeatability of proofs from a limited environmental situation, such as a herd, a non-linear increase of daughters over herdmates, in different genetic levels, could be suspected.
2. It is possible that there could be a genotype by environmental interaction with sire's genotypes interacting with herd environments.
3. Another possible cause, for lack of repeatability, is environmental correlations among progeny and/or herdmate groups.

Enough evidence is available to indicate that there is little or no possibility for the regression of genotypic level on increase of daughters over herdmates, to be non-linear. The regression of daughter on dam is almost completely linear over a wide range of production. Paternal half-sib estimates of heritability have been observed to increase as level of production increases; if not markedly, at least in the same direction. Heritability of milk production seems to be nearly linear over the range in the current population. Sire by herd-year-season interactions, if they are of considerable magnitude, could conceivably present some problems in sire evaluation. Several studies have pointed out that this interaction could be a source of error for bulls that are used in limited number of herd-year-seasons, but not likely so when bulls are used extensively over many herd-year-seasons. In general, accuracy of proofs increases as the bulls are tested in more herd-year-seasons with a fairly large number of stablemate sires, and essentially small number of daughters per stablemate sire.

The problem of environmental correlation among relative groups seems important and needs investigation. From the available evidence it appears that this is one reason, probably a major one, why the initial proof that is made in a limited environmental situation is not highly repeatable in AI. The daughter average is not a good measurement of a sire's genetic merit when all the daughters are in the same herd, that is, natural proof situation. The possibilities of bias are too great, and the reliability of the estimate is low since the variance of the average is inflated by the environmental covariances among daughters. These covariances include all of the herd effect for all daughters, and also

time effects for cows calving in the same year-season. Under these conditions, increasing the number of daughters does not greatly reduce the biases or increase the value of the daughter average. But, if the environmental correlation between the daughters is considered, along with the distribution of daughters over herds, a better estimate of the sire's genetic merit is expected.

Another indication that points in the direction of environmental correlation among paternal half-sibs contributing to inaccurate sire evaluation, is in the fact that any sort of estimates, based on half-sib components, are not very meaningful unless they have been made from sires who have daughters in several herds. While theoretically, deviation records are designed to eliminate all environmental effects related to herd and year-season of freshening, this actually is not the case in the average run of data. Specifically, when more than one daughter of a single bull is involved in a particular herd-year-season group, then some correlation does remain between the deviation records.

The present study was initiated to explore further the importance of such environmental correlations. Environmental correlation among paternal half-sisters was investigated since they form the largest group of relatives in the current dairy cattle population. Also, the sire selection is the most effective means of making genetic change in the population. This environmental correlation, often represented by c^2, is the extra correlation among paternal half-sisters, contributed by factors other than their sire alone, which paternal sisters may have in common but which may differ from one set of paternal sisters to another. For example, a bull's daughters calving in the same herd-year-season are

expected to be more closely correlated than daughters calving in the same herd but in different year-seasons.

II. REVIEW OF LITERATURE

In the early thirties, Lush (1931 and 1935), using Wright's (1921 and 1934) path coefficient method, theoretically demonstrated various aspects of environmental correlation existing between different relative groups. He shows that the limiting value of the correlation between the true merit of a bull and the average of single records of his daughters is reached, when the number of daughters becomes indefinitely large and depends upon heritability and the environmental correlation among the daughters. When the environmental correlation is zero and heritability is non-zero, the correlation has a limit of 1.0 regardless of the magnitude of heritability. However, when the environmental correlation is present, this limit is always less than one, and may be considerably lower, the exact magnitude depending upon the magnitude of the two parameters. Increasing the number of daughters per bull tends to eliminate errors arising from random uncorrelated variations in environments but does not tend to eliminate errors arising from dominance, or epistasis or consistently biased and uncorrected environmental effects.

Lush and McGilliard (1955) discussed some of the possible sources of bias that may enter progeny tests. This bias is the amount by which the daughters of a sire are above or below the breed average because of circumstances, other than the sire's breeding value, which tend to be alike for all his daughters. The largest part of the bias comes from the environmental conditions which are similar for all daughters of the same sire but vary from one sire to another. Part of the differences caused by dominance, or by epistasis, or by interactions between heredity

and environments will contribute to the bias as will the similarities in the breeding values of the mates of a sire. Johansson (1961) in his book on dairy cattle breeding, discusses this paper and suggests that the systematic differences between the progeny groups of a non-genetic nature arising from feeding and management, season of calving, age at calving, etc., contribute to the phenotypic correlation between members of the same progeny group.

Johansson (1960) studied repeatability of the progeny test as a measure of accuracy where he discussed the regression of future daughters on those tested. If the various components of variance between progeny groups are represented by G_s, for genetic variance due to differences between sires; C_s, for environmental variance between progeny groups; and R, for variance due to randomly distributed causes, then the regression of future daughters on those tested, that is, the repeatability of the progeny test, is:

$$b = \frac{G_s}{G_s + C_s + R/n} = \frac{0.25\ h^2 n}{1 + (n-1)(0.25\ h^2 + c^2)} \qquad 1$$

where n is the number of daughters, h^2 is the heritability of the trait, c^2 is the fraction of variance which is due to non-genetic differences between groups and 0.25 is the coefficient of relationship between the paternal half-sisters (random mating). The correlation between paternal half-sisters arising from non-genetic source is $c^2 = \frac{C_s}{G_s + C_s + R/n}$ while the phenotypic correlation between them is $t = 0.25\ h^2 + c^2$.

The assumption here is that there may be an environmental correlation, c^2, between the tested daughters of the same bull, but no such correlation

exists between the tested daughters and future daughters, as would be true if each bull is tested in one herd and used later in any other herd of that breed.

The regression increases with increasing number of daughters, but it decreases with rising values of c^2 as shown in the Figure 1.

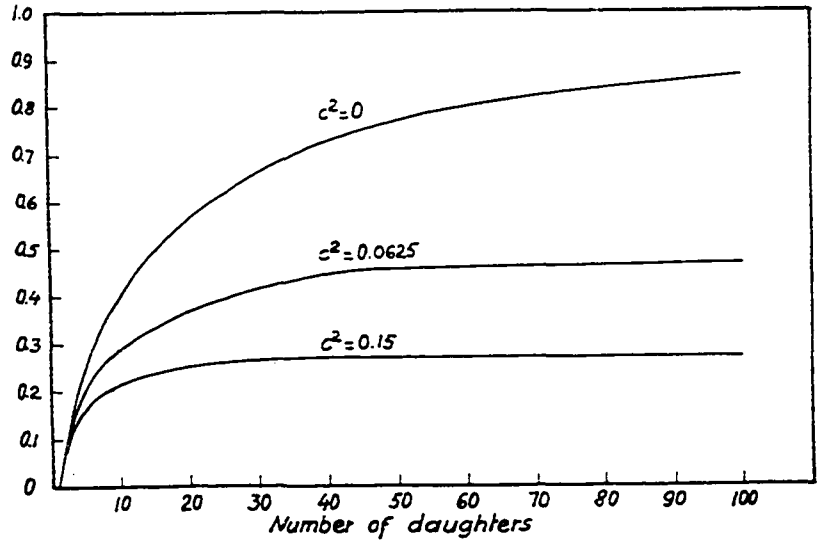

Figure 1. The repeatability of the progeny test based on varying number of daughters (n) when there is no environmental correlation between tested daughters and future daughters, but under three different assumptions with regard to the environmental correlation within the progeny groups: (1) $c^2 = 0$; (2) $c^2 = 0.0625$ and (3) $c^2 = 0.15$. The heritability is assumed to be 0.25 in all cases (Johansson 1961, p. 222).

Gaunt and Legates (1958) considered the number of daughters and number of herds to be important for estimating regression of sire's merit on the daughter average:

$$b = \frac{2\,\sigma_s^2}{\sigma_s^2 + \frac{\sigma_k^2}{n_h} + \frac{\sigma_e^2}{n_d}} \qquad 2$$

where n_h and n_d are number of herds and daughters, respectively. This implies that the number of daughters per herd is constant for all herds. However, it does consider the possibility that a sire may have several daughters per herd; consequently, the increase in the variance of daughter average caused by the environmental correlation among daughters in the same herd, is taken into account.

Touchberry, Rottensten and Andersen (1960) analyzed the first lactation milk and milk fat records of daughters of Red Danish Milkrace sires tested at Danish bull testing stations. They report that the expected correlation between the averages of the progeny group at the test stations X and those averages from field results as measured by Y, can be expressed as:

$$r_{XY} = \frac{g_X}{2}\sqrt{\frac{N_X}{1 + (N_X-1)t_X}}\; r_{G_XG_Y}\; \frac{g_Y}{2}\sqrt{\frac{N_Y}{1 + (N_Y-1)t_Y}} \qquad 3$$

where g_X and g_Y are the square roots of the heritabilities obtained at the test stations and former herds, respectively. N_X and N_Y are the corresponding average number of daughters per sire and t_X and t_Y are the corresponding phenotypic correlations between daughters of a sire, which can be expressed as $t_X = 1/4\ g_X^2 + c_X$ and $t_Y = 1/4\ g_Y^2 + c_Y$, c_X and c_Y being the corresponding environmental correlations, respectively. They estimated environmental correlation, in test stations, to be 0.09 and 0.07 for milk and milk fat, respectively.

Heidhues, VanVleck and Henderson (1961) studied the prediction of future daughter's production based on 10 successive groups of 10 daughters each and cumulative groups of the first 10, 20, . . . 100

daughters for 53 Holstein-Friesian sires used in artificial breeding in New York. The production of 200 daughters that made records after the first 100 daughters, was taken to be a measure of the true breeding value of the sire. The expression to estimate the regression of the mean of q daughters on the production of future daughters, was q / (q + 17). The correlations based on successive groups of 10 daughters and 200 future daughters were slightly lower than expected for milk, those of fat being almost equal to expected values. There was no significant difference to indicate presence of any environmental correlation.

A similar study was made by McDaniel and Corley (1965) using records of AI progeny of 277 bulls with 240 or more first lactation records each. The first 120 progeny were divided into 12 sets of 10 daughters each; the 13^{th} set of 121 - 240 was used to determine the actual breeding value of the bull. Their results are very similar to those formed by Heidhues et al. (1961).

Meek and Van Vleck (1964) studied the correlations between: 1) non-AI daughter-dam comparison, 2) AI daughter-dam comparison, 3) non-AI herdmate comparison, 4) USDA AI herdmate difference, and 5) Cornell daughter level. From the low correlations, between natural service and AI service evaluations, the authors conclude that environmental correlations in natural service proofs are important and should be considered in the regression of the sire's true genetic value on the daughter average.

Bereskin and Lush (1965) discuss the estimation of correlation ($r_{G_S\bar{D}}$) between the genotype of a bull (G_S) and his daughter average ($\bar{D}$), which is:

$$r_{G_S\bar{D}} = \frac{h}{2}\left[\frac{n}{1 + (n-1)\, r_{D_iD_j}}\right]^{1/2} \qquad 4$$

where $r_{D_iD_j} = 1/4\ h^2 + c^2$.

Prominent possible causes of c^2 are correlated environmental effects, correlations between the breeding values of the mates of the sire, correlations between the breeding values of the sire and his mates, and correlation involving both environmental and genetic effects. The correlation in equation 4 squared is the regression which is the same as in equation 1. The equation predicting the average production of m future daughters ($\bar{D}_m$) from the average production of n tested daughters ($\bar{D}_n$), as deviated from population mean (μ) is:

$$\hat{\bar{D}}_m = \frac{nw}{1 + (n-1)t}\,(\bar{D}_n - \mu) \qquad 5$$

where t represents $r_{D_iD_j}$ among n tested daughters, and w is like t except that w allows for the possibility that c^2 may differ in two daughter groups i.e. $w = h^2 / 4 + c^2_{nm}$, where c^2_{nm} may or may not be as large as c^2.

If production records are expressed as deviations from their respective contemporary herdmate averages (HA), then

$$r_{G_S(\overline{D-HA})} = \left[\frac{nkh^2}{4\,(1-s)(k-1) + (n-1)(h^2+4c^2)k}\right]^{1/2} \qquad 6$$

where k is the number of contemporary herdmates; s is the average correlation among the records of the cow, including the daughter, in

the same herd-year-season group, each with but one record, in the proof. A theoretical example is given in the Table 1 from Bereskin and Lush (1965).

Table 1. Expected values for $r_{G_S(D-HA)}$ with $h^2 = .25$, $s = .32^a$, $k = 10$, and $c^2 = .00$, .01, .02, .04, and .06

n	c^2				
	.00	.01	.02	.04	.06
5	.56	.55	.54	.52	.50
10	.69	.67	.65	.61	.58
20	.80	.77	.73	.69	.64
30	.86	.81	.77	.71	.66
50	.91	.85	.81	.74	.68
70	.93	.87	.82	.75	.69
100	.95	.89	.84	.76	.70
00	1.00	.93	.87	.78	.71

[a]s represents the estimated portion of the total variance in production records directly attributable to herd-year-season effects.

The authors computed the actual correlations for first available record and the average of all records, for both deviated and non-deviated records, for each daughter and compared them with the expected correlations based on number and kind of records included. When the distribution of records was similar to that existing in idealized AI situation, the computed correlations were significantly smaller than the expected correlations. They concluded that the use of deviated records only slightly reduced the difference between the correlation. The differences were due, at least in part, to the presence of residual correlation between paternal half-sisters (c^2). Factors contributing to c^2 should be given fully as much consideration as numbers of daughters per se in bull proofs,

a major factor being the distribution of daughters among herd-year-seasons.

Allaire and Gaunt (1965) studied the contemporary records to see if a measure comparising a first lactation record with its first lactation contemporary records, is superior to using the average of all contemporary records, and found that all lactation contemporaries overestimated the environmental situation for first lactation records by 285 pounds of milk.

Van Vleck and Hart (1965) discussed daughter-dam regression (b_{dd}), full-sib correlation (r_{fs}) and maternal half-sib correlation (r_{ms}), in the same and different herds and formed the following equations:

$$2b_{dd} = A_o^2 + 2\ 1/2\ A_oA_m + A_m^2 + 1/2\ A_{oo}^2 + 2C_{dd}^2$$

$$2r_{fs} = A_o^2 + 2A_oA_m + 2A_m^2 + 1/2\ A_{oo}^2 + 1/2\ D_o^2 + 2C_{fs}^2$$

$$4r_{ms} = A_o^2 + 2A_1A_m + 4A_m^2 + 1/4\ A_{oo}^2 + 4C_{ms}^2$$

where A_o^2 is additive genetic variance; A_oA_m is the covariance between maternal additive genetic and additive genetic effects; A_{oo}^2 is additive-by-additive genetic variance; D_o^2 is dominance genetic variance; and C^2 is environmental covariance (dd between daughter and dam, fs between full-sibs and ms between maternal half-sibs in the same herd). When the pair of records are from different herds the C^2 term drops out. The authors conclude that the environmental covariance between first lactation deviation records of: 1) daughters and dams in the same herd is rather small, .01; 2) maternal half-sibs in the same herd is small, .00; 3) full-sibs in the same herd may be important and probably is between .06 and .12 of the total variance.

Using the first lactation deviations from the herd averages,

Van Vleck (1966b) compared the actual and expected correlations between groups of AI sired daughters of the same Holstein bulls separated by various time intervals. The correlations were computed in three different ways: 1) within groups, 2) between groups separated by 8, 12, 24, 36 and 48 months, and 3) considering differing numbers of daughters in each group. The expression used to compute the actual correlation is $(s_{12} + c_{12}) / [(s_1 + c_1 + e_1 / n_1) (s_2 + c_2 + e_2 / n_2)]^{1/2}$ which is similar to one given by Bereskin and Lush (1965); the expected correlations were computed as $[(n_1 / (n_1 + e_1 / s_1)) (n_2 / (n_2 + e_2 / s_2))]^{1/2}$; where s_{12} is the genetic covariance between paternal half-sibs in the two groups, due to their having the same sire; c_{12} is the environmental covariance between the two groups; c_1 and c_2 are the environmental covariances among half-sibs in group 1 and group 2, n_1 and n_2 are the number of records in the respective groups, e_1 and e_2 are the corresponding estimates of the within sire components of variance, and s_1 and s_2 are the estimated sire components of variance, respectively.

There was no apparent pattern to the differences in correlations found, either over time or by sire of daughter group studied. Van Vleck concluded that environmental correlations are small or nil among AI sired half-sisters, and are unimportant as a source of error in evaluation of sires in New York State. However, the evidence present does not preclude a small environmental correlation masked by sampling error and the evidence is restricted to AI sired daughters.

Van Vleck (1966c) made a similar study with first lactation records of paternal half-sisters. Three groups of data were used, daughters of natural service sires, AI sires and all sires regardless of type of

service. Each of these was subdivided into sets of data consisting of: 1) a set of three daughters of each sire, two in one herd and the third in a different herd, or 2) a set of two daughters of each sire, both in same herd, or 3) a set of two daughters of each sire in different herds. A sire could be represented in only one of the above three sets.

Comparison of the correlation among half-sisters in the same and different herds was used to estimate the magnitude of the environmental correlation, c^2, between paternal half-sibs. Such correlations were estimated to be .168 and .086 when the pairs were in same and different herds, respectively, suggesting the magnitude of c^2 to be about .082 of the total variance. Van Vleck found no evidence to indicate that the environmental correlation is lower for AI sired pairs than for naturally sired pairs in the same herd, rather the reverse was observed, .088 vs. .130. No important environmental correlation among maternal half-sisters in the same herd was observed. He concluded that c^2 is relatively large and certainly should be considered in natural service sire evaluation.

McDaniel[1] set up a series of 27 equations that were formed by regressing progeny groups, differing in number of herds and distributions of daughters over herds, on a very large independent sample of progeny. The equations are:

$$k_i h^2 + m_i c^2 = b_i$$

where k_i are coefficient of heritability (h^2) that depend on the number of daughters in the sample, m_i are coefficients of environmental correlation

---

[1]B. T. McDaniel. 1967 and 1968. U.S. Department of Agriculture, Beltsville, Maryland. Simultaneous estimation of h^2 and c^2. Private communication.

(c^2) that depend on 1) the number of herds represented in the sample of progeny and 2) the distribution of progeny over these herds, and b_i are the regressions.

These equations were solved by using weighted least squares and the simultaneous solutions were obtained for h^2 and c^2 which were 0.19 and 0.14, respectively.

Plowman and McDaniel (1968) in a discussion of the methods currently being used to estimate breeding value of bulls by the U.S. Department of Agriculture, show how the presence of a residual correlation (c^2) among paternal half-sibs in the same herd effects the regression of future progeny on initial progeny. They suggest this regression to be:

$$\frac{N\,h^2}{4 + (N-1)\,h^2 + 4\,\frac{\Sigma n_i(n_i-1)}{N}\,c^2} \qquad 7$$

where n_i is the number of progeny in i^{th} herd, N is the total number of progeny of the bull, and c^2 is the residual correlation among half-sibs in the same herd after they are expressed as deviations from herdmates.

If a daughter has more than one record they suggest replacing N by ΣW_j where W_j is defined as:

$$W_j = \frac{n_j R}{R[1 + (n_j-1)R]}\ ; \qquad 8$$

R is the repeatability of individual records and is taken as 0.50, and n_j is the number of records of the j^{th} daughter. This expression is the regression for estimating future production of a cow from past records coded by R in the denominator so $W_j = 1$ for one record as a base. This

was thought to be more understandable by dairymen, since this is a practical tool which dairymen use.

Thus, in addition to utilizing information on number of herds and distribution of daughters across herds, this procedure accounts for the number of records per daughter as well. Using values of 0.19 for heritability of single records and 0.14 for the residual correlation among half-sibs they demonstrated the effect of number of daughters, number of herds, and the distribution of daughters among herds as shown in Figure 2.

Regression of future progeny on initial progeny as affected by number of daughters, number of herds, and the distribution of daughters among herds—$h^2 = .19$[a]; $c^2 = .14$

Number of daughters	1 daughter per herd		5 daughters per herd[b]		10 daughters per herd[b]		All daughters in a single herd	
	b	No. of herds	b	No. of herds	b	No. of herds	b	No. of herds
5	.200	5	.136	1			.136	1
10	.333	10	.239	2	.177	1	.177	1
15	.429	15	.321	3			.197	1
20	.500	20	.386	4	.301	2	.209	1
25	.556	25	.440	5			.216	1
30	.600	30	.486	6	.392	3	.221	1
40	.667	40	.557	8	.463	4	.229	1
50	.714	50	.612	10	.518	5	.233	1
70	.778	70	.688	14	.601	7	.239	1
100	.833	100	.759	20	.683	10	.243	1
200	.909	200	.863	40	.812	20	.248	1

[a] If each daughter is in a different herd, this gives the same regression value as the equation $\frac{n}{n+20}$ now in use for AI sire summaries.

[b] Equal number of daughters in each herd.

Figure 2. A theoretical example of regressions (Plowman and McDaniel 1968, p. 309).

Thomson (1968) studied the environmental correlations between paternal half-sisters. Records of half-sibs were subdivided into four categories: 1) same herd and same year-season, 2) same herd but different year-seasons, 3) different herds but same year-season, and 4) different herds and different year-seasons. The corresponding components of

variance for paternal half-sibs are $(\sigma_s^2 + \sigma_a^2 + \sigma_h^2 + \rho^2\sigma_e^2)$, $(\sigma_s^2 + \sigma_h^2 + \rho^2\sigma_e^2)$, $(\sigma_s^2 + \sigma_h^2)$ and σ_s^2 respectively, where σ_s^2, σ_a^2, σ_h^2 and σ_e^2 are components of variance due to sire, year-season, herd and error, and ρ^2 is the additional environmental correlation between half sibs calving in the same herd.

The environmental correlations s, t and w were estimated as the intra-class correlations derived from between and within herd-year-season, herd and year-season analysis respectively, for both mature-equivalent and deviation records. Using all records in analyses for correlation (s) among cows in the same herd and same year-season, but only one cow's record selected at random from each herd-year-season for correlation (t) among cows in same herd but different year-seasons and the correlation (w) among cows in same year-season but different herds, the estimates were 0.328, 0.226 and 0.028 respectively, for ME records and 0.0, 0.0 and 0.003 respectively, for deviation records. The estimate for environmental correlation (ρ^2) as a measure of common environments that half-sibs in the same herd experience, over and above the correlation due to a common sire and common herd effect, was 0.005 for ME records and 0.102 for deviation records.

From the previous discussion it is evident that some environmental correlation (c^2) does exist among relative groups in the same herd. Its magnitude varies from one report to another, depending on the relative groups involved and the method of estimation used. The correlation appears to be small among relative groups such as daughter-dam or maternal half-sibs, but is considerably large among relatives like full-sibs or paternal half-sibs (Van Vleck 1966b and 1966c, Touchberry et al. 1960,

McDaniel[1] 1967 and 1968 and Thomson 1968). Different methods of estimating environmental correlations do not actually estimate the same parameter. There has been no consistent definition of c^2 largely because it has been estimated differently for different applications. The present work was under taken to study some of the various forms of c^2 and pinpoint those that are important from practical point of view.

---

[1] B. T. McDaniel. 1967 and 1968. U.S. Department of Agriculture, Beltsville, Maryland. Simultaneous estimation of h^2 and c^2. Private communication.

III. SOURCE AND ADJUSTMENT OF DATA

Records of grade and registered Holsteins in Official DHI and DHIR herds (active herds only) in eight midwestern states were obtained from the Iowa Dairy Record Processing Center at Iowa State University. The states were Iowa, Missouri, North Dakota, South Dakota, Nebraska, Kansas, Arkansas, and Oklahoma. The following requirements were set for a record to be used for the present study.

1. The lactation must have begun during the years 1957 through 1967.
2. Only first lactation records were included. A lactation that began at the age of 22 through 35 months was considered as first lactation record. All available lactation records were used to compute herdmate averages.
3. A record was discarded if -
   a) the sire of the cow was not known,
   b) the lactation began without the first monthly report following calving,
   c) the record was estimated for two or more consecutive months during the lactation,
   d) the lactation began with, or was terminated by an abortion prior to 305 days in milk,
   e) the lactation was terminated by sale or death of the cow prior to 305 days in milk,
   f) the cow was milked three times per day,
   g) the cow was used as nurse cow,

h) the herdmate average production of milk and butterfat were not available.

All completed lactations and all lactations in progress 45 days or more were used. The herdmate average production of a cow was obtained as the average production of all other cows of her breed, in the herd, that freshen during the same year-season. All records were standardized to a 305 day, twice-a-day milking, mature-equivalent basis. They were used to calculate herdmate averages. These averages have been calculated by the Processing Center and were used as such. Similarly, mature-equivalent (ME) production was used as provided, with each lactation, by the Processing Center. The ME factors that had been used were those by Kendrick (1955).

The restrictions on the data should leave only records which represent normal lactations that are unaffected by severe conditions. A cow that was turned dry for any reason prior to reaching 305 days in milk, was included in the study. The number of sires was 1963 with a total of 55,170 first lactation records made in 2,326 herds. The units digit was dropped from the milk production records, for all the analysis of variance.

The lactations were divided into two seasonal groups as recommended by Bereskin and Freeman (1965); that is May through September and October through the following April. They used these season initially with Iowa data and determined them as the most appropriate. Freeman[1] evaluated this

---

[1]A. E. Freeman. 1969. Iowa State University of Science and Technology, Ames, Iowa. Comments on estimation of environmental correlations. Private communication.

seasonal grouping several times later and found that the seasons of May - September and October - April seem to be the best. Thus, there were 23 year-seasons in the data. The first year-season included records initiated between January 1957 and April 1957.

IV. DERIVATION OF REGRESSIONS

Since the purpose here is to estimate the term c^2 - the additional environmental correlation arising among daughters of a bull which calved in the same herd and/or year-season, as compared to those calving in different herds and/or year-seasons - the regressions or correlations that involve c^2 can be used as a basis for estimation. Some basic assumptions that were made for all estimates of c^2 were:

1. Only one record per daughter.
2. No correlation between the breeding value of a sire with:
   (a) the genetic merit of his mates,
   (b) the herd effects in the herds in which he was used,
   (c) the permanent environmental factors affecting daughter's production.
3. Environmental covariances among daughters calving in the same herd are similar, regardless of the year of calving.
4. No environmental correlations between progeny in one herd and progeny in another.
5. Expectations of variances of environmental effects are similar from herd to herd.

A lactation record was represented by the following model:

$$Y_{ijk} = \mu + h_i + s_j + e_{ijk}$$

where,

Y_{ijk} is the record of k^{th} daughter by the j^{th} sire made in the i^{th} herd,

μ is the population mean,

h_i is an effect common to all daughters in the i^{th} herd,

s_j is an effect common to all daughters of the j^{th} sire, and

e_{ijk} is an effect peculiar to record of the k^{th} daughter by the j^{th} sire made in the i^{th} herd, $(k = 1, 2, \ldots, n_{ij})$.

A. Regression of the Average of Daughters in a Sample on the Average of Daughters in Another Sample

The daughters of a bull were randomly divided into two samples, S_M and S_N. For this section, no restriction about the distribution of daughters in herds among the two samples was made. Let Y_{ij} represent the j^{th} daughter in the i^{th} herd in sample S_M and Y'_{il} represent the l^{th} daughter in the i^{th} herd in sample S_N, where $i = 1, 2, \ldots, k$; $j = 1, 2, \ldots, m_k$ and $l = 1, 2, \ldots, n_k$; also $\Sigma m_i = M$ and $\Sigma n_i = N$, where m_i is the number of daughters in the i^{th} herd in sample S_M and n_i is the number of daughters in the i^{th} herd in sample S_N.

The averages of daughters in the two samples are:

$$\bar{Y}_M = \frac{Y_{11} + Y_{12} + \ldots + Y_{21} + \ldots + Y_{km_k}}{M}$$

$$\bar{Y}_N = \frac{Y'_{11} + Y'_{12} + \ldots + Y'_{21} + \ldots + Y'_{kn_k}}{N}$$

and the covariance:

$$\text{Cov}\,(\bar{Y}_M, \bar{Y}_N) = \frac{1}{MN}\left[\Sigma m_i n_i\, \sigma_{y_{ij}y'_{il}} + (MN - \Sigma m_i n_i)\, \sigma_{y_{ij}y'_{i'l}}\right] \qquad 9$$

where $\sigma_{y_{ij}y'_{il}}$ is the covariance between daughters of the two samples in the same herd, and

$\sigma_{y_{ij}y'_{i'1}}$ is the covariance between daughters of the two samples in different herds;

which by definition are:

$$\sigma_{y_{ij}y'_{il}} = \sigma_s^2 + e^2 \qquad 10$$

and

$$\sigma_{y_{ij}y'_{i'1}} = \sigma_s^2$$

where σ_s^2 is the sire component of variance and e^2 is the additional effect due to daughters being in the same herd.

Substituting the values from equation 9 in equation 10:

$$\text{Cov } (\bar{Y}_M, \bar{Y}_N) = \frac{1}{MN} [\Sigma m_i n_i \ (\sigma_s^2 + e^2) + (MN - \Sigma m_i n_i) \ \sigma_s^2]$$

which reduces to:

$$\text{Cov } (\bar{Y}_M, \bar{Y}_N) = \sigma_s^2 + \frac{\Sigma m_i n_i}{MN} e^2 \ . \qquad 11$$

The variance of the average of daughters in sample S_M is:

$$V(\bar{Y}_M) = \frac{1}{M^2} [M \sigma_y^2 + \Sigma m_i (m_i - 1) \ \sigma_{y_{ij}y'_{il}} + \sum_{i \neq i'} m_i m_{i'} \ \sigma_{y_{ij}y'_{i'1}}]$$

where $\sigma_{y_{ij}y'_{il}}$ and $\sigma_{y_{ij}y'_{i'1}}$ are defined earlier in 10. Substituting their values:

$$V(\bar{Y}_M) = \frac{1}{M^2} [M \sigma_y^2 + \Sigma m_i (m_i - 1) \ (\sigma_s^2 + e^2) + \sum_{i \neq i'} m_i m_{i'} \ \sigma_s^2]$$

which reduces to:

$$V(\bar{Y}_M) = \frac{1}{M}\left[\sigma_y^2 + (M-1)\ \sigma_s^2 + \frac{\Sigma m_i(m_i-1)}{M}\ e^2\right] \qquad 12$$

By symmetry the variance of average of daughters in sample S_N is:

$$V(\bar{Y}_N) = \frac{1}{N}\left[\sigma_y^2 + (N-1)\ \sigma_s^2 + \frac{\Sigma n_i(n_i-1)}{N}\ e^2\right] \qquad 13$$

The regression of $\bar{Y}_N$ on $\bar{Y}_M$ is the covariance term divided by $V(\bar{Y}_M)$, i.e. $b_{\bar{Y}_N \bar{Y}_M}$; denoting it by b_1:

$$b_1 = \frac{M\ \sigma_s^2 + \frac{\Sigma m_i n_i}{N}\ e^2}{\sigma_y^2 + (M-1)\ \sigma_s^2 + \frac{\Sigma m_i(m_i-1)}{M}\ e^2}$$

Dividing both numerator and denominator by $4\ /\ \sigma_y^2$:

$$b_1 = \frac{M\ h^2 + 4\ \frac{\Sigma m_i n_i}{N}\ c^2}{4 + (M-1)\ h^2 + 4\ \frac{\Sigma m_i(m_i-1)}{M}\ c^2} \qquad 14$$

where,

$h^2 = 4\ \sigma_s^2\ /\ \sigma_y^2$ = heritability and,

$c^2 = e^2\ /\ \sigma_y^2$ = the additional environmental correlation.

B. Regression of Breeding Value on the Daughter Average

If all the daughters in a herd appear in one and only one of the two samples, S_M or S_N, the covariance term in regression b_1 contains only the sire component of variance, since the other term just counts the

members of each sample in a common herd, and the expression simplifies to:

$$b = \frac{M\,\sigma_s^2}{\sigma_y^2 + (M-1)\,\sigma_s^2 + \frac{\Sigma m_i(m_i-1)}{M}\,e^2}$$

$$= \frac{M\,h^2}{4 + (M-1)\,h^2 + 4\,\frac{\Sigma m_i(m_i-1)}{M}\,c^2} \qquad 15$$

where b is the regression of breeding value on the daughter average.

This expression for the regression can also be obtained from the path diagram in Figure 3:

$$1 = M\,p^2 + M\,(M-1)\,p^2\,(\tfrac{h^2}{4}) + [m_1(m_1-1) + m_2(m_2-1) + \ldots + m_1(m_1-1)]\,c^2p^2$$

$$= M\,p^2 + M\,(M-1)\,p^2\,(\tfrac{h^2}{4}) + \Sigma m_i(m_i-1)\,c^2p^2$$

$$= p^2\,[M + M\,(M-1)\,\frac{h^2}{4} + \Sigma m_i(m_i-1)\,c^2]$$

Therefore:

$$p^2 = \frac{1}{M[1 + (M-1)\,\frac{h^2}{4}] + \Sigma m_i(m_i-1)\,c^2}$$

The correlation of the breeding value with daughters' average is:

$$r_{G_s \cdot \bar{D}} = \frac{1}{2}\,Mhp$$

$$= \frac{1}{2}\,Mh\,\sqrt{\frac{1}{M[1 + (M-1)\,\frac{h^2}{4}] + \Sigma m_i(m_i-1)\,c^2}}$$

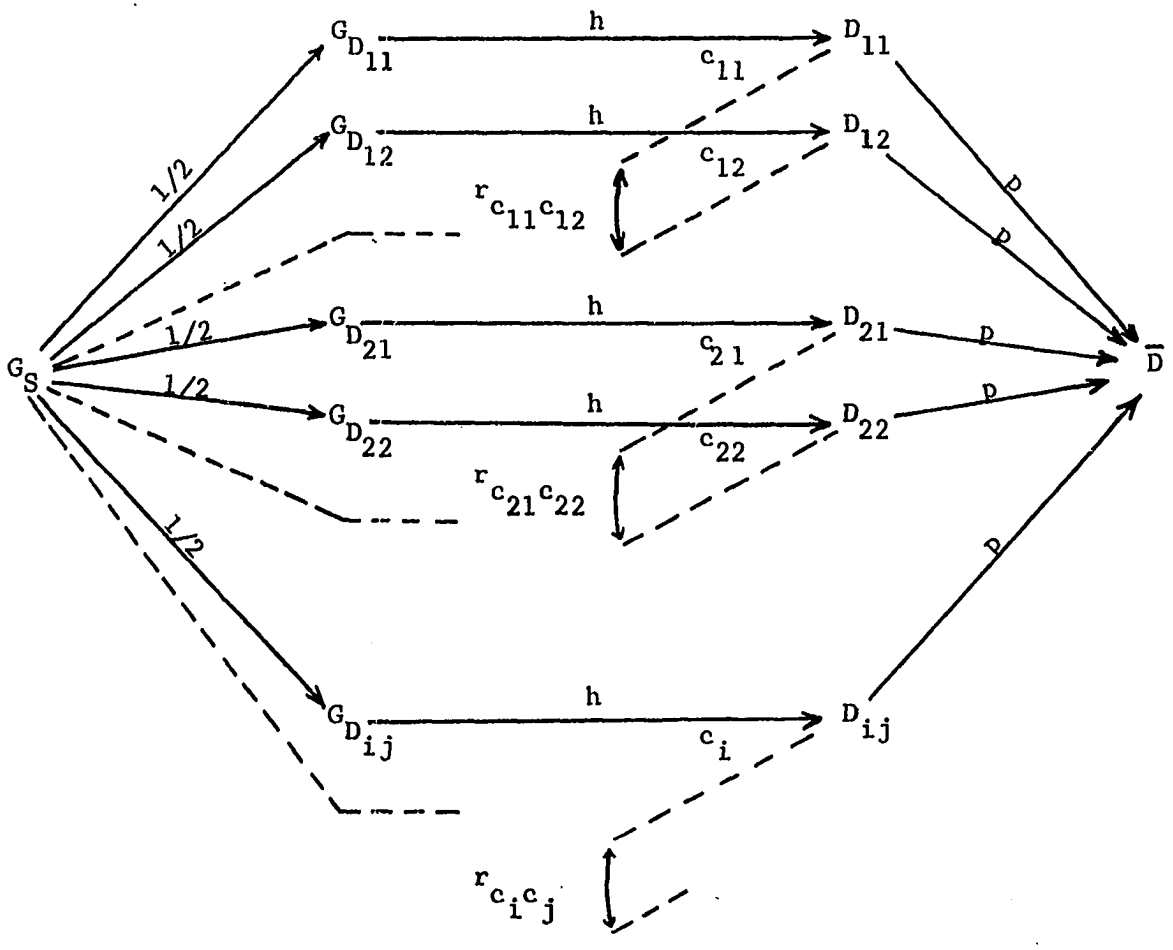

Figure 3. Path diagram illustrating the correlation of breeding value on the daughter average.

The regression of breeding value on daughters' average is simply the square of the correlation. Therefore:

$$b_{G_s \cdot \bar{P}} = \frac{M\, h^2}{4 + (M-1)\, h^2 + 4\, \dfrac{\Sigma m_i (m_i - 1)}{M}\, c^2} = b \ .$$

If all the daughters are in one herd, $\Sigma m_i = M$ and the coefficient of c^2 is 4 M (M-1); but if each daughter is in a different herd then $m_i = 1$ for all i's and the coefficient of c^2 is zero.

V. RESULTS AND DISCUSSION

A. Heritability

Heritability was estimated from a between and within sire analysis of variance using both mature equivalent and deviation from herdmate average data. The estimates of h^2 were computed in the usual way so as to compare them later with the values obtained by simultaneous estimation of h^2 and c^2; thereby, giving a basis for placing some confidence on the c^2 estimates. The following model was assumed:

$$Y_{ij} = \mu + s_i + e_{ij}$$

where,

Y_{ij} is the first lactation record of the j^{th} daughter of the i^{th} sire,

μ is the population mean,

s_i is an effect common to all daughters of the i^{th} sire, and

e_{ij} is a random effect associated with the j^{th} daughter of the i^{th} sire.

Estimates of heritability were obtained from the following analysis of variance:

Source of Variation	d.f.	E.M.S.
Between sires	s-1	$\sigma_e^2 + \frac{N}{s-1}\left(N - \frac{\sum_i n_i^2}{N}\right)\sigma_s^2$
Within sires	N-s	σ_e^2
Total	N-1	

where heritability is estimated as four times the intra-class correlation:

$$h^2 = \frac{4\ \sigma_s^2}{\sigma_s^2 + \sigma_e^2} .$$

The variance of the intra-class correlations were calculated by the following approximation of Swiger, et al. (1964):

$$V(\hat{t}) \simeq \frac{2\ (N-1)(1-t)^2\ [1 + (k-1)t]^2}{k^2\ (N-s)(s-1)}$$

where s is the number of sires, t is the intra-class correlation, N is the total number of daughters, k is the coefficient of the sire component of variance, i.e.:

$$k = \frac{1}{s-1} \left(N - \frac{\Sigma\ n_i^2}{N}\right) .$$

Since the variance of the estimate of heritability is sixteen times the variance of the intra-class correlation, the standard error of the estimate of heritability is:

$$\sigma_{(\hat{h}^2)} = 4 \quad V(\hat{t})$$

Sire groups that had less than four daughters or daughters only in one herd were not included in the analyses for heritability estimation. Only one daughter per herd was used (selected at random) and only first lactation records were included. Since some of the records were not numbered by lactations, a record was considered to be first lactation record if the cow was 22 to 35 months of age at the time of first calving. Obviously, cows that freshened at less than 22 months of age or that did not have a record with a freshening date within 22 to 35

months of age were excluded from the analyses.

Table 2 contains analyses for both mature-equivalent and deviation records for the two traits, milk and milk fat.

The heritability estimates of milk and milk fat for mature-equivalent and deviation records were 0.257, 0.200, 0.211 and 0.163, respectively. The estimates for milk are in agreement with those generally found in literature; those for milk fat though, are slightly lower. Blanchard, et al. (1966), using Iowa DHIA Holstein data obtained estimates of the heritability of deviation milk and milk fat to be same, 0.29. He obtained these estimates from paternal half-sib correlations with the restriction that each sire had at least three daughters. Van Vleck and Brandford (1965) estimated heritability at 0.24 for deviation milk. They used about 20,850 first lactation records of Holstein in New York State collected from 1950 through 1963.

Butcher (1965) estimated the heritability of first lactation milk yield from paternal half-sib correlations using Iowa DHIA data and California data. He obtained values of 0.279 and 0.382 for the two sets of data, respectively. When the Iowa data were restricted to the daughters of a sire used in at least four herds, the estimate for first lactation deviation records increased to 0.377.

Thomson (1968) computed heritability estimates for both mature-equivalent and deviation records, which were 0.338 and 0.358. These values were estimated using field data from all registered Holstein cows. The cows were located in herds that had 'proven' a bull. It is possible that these herds were managed nearly alike. Therefore, the estimates for heritability are likely to be somewhat higher due to the relatively

Table 2. Analysis of variance - heritability estimates

MATURE-EQUIVALENT RECORDS

Source of variation	d.f.	M.S.	E.M.S.
Milk:			
Between sires	812	223,082.9	$\sigma_e^2 + 24.62\ \sigma_s^2$
Within sires	19,266	82,961.8	σ_e^2
	$\hat{\sigma}_e^2 = 82{,}961.8$	$\hat{\sigma}_s^2 = 5{,}690.4$	$\hat{h}^2 = 0.257 \pm 0.019$
Milk fat:			
Between sires	812	25,115.1	$\sigma_e^2 + 24.62\ \sigma_s^2$
Within sires	19,266	10,950.7	σ_e^2
	$\hat{\sigma}_e^2 = 10{,}950.7$	$\hat{\sigma}_s^2 = 575.2$	$\hat{h}^2 = 0.200 \pm 0.017$

DEVIATION RECORDS

Source of variation	d.f.	M.S.	E.M.S.
Milk:			
Between sires	812	162,639.1	$\sigma_e^2 + 24.62\ \sigma_s^2$
Within sires	19,266	68,519.9	σ_e^2
	$\hat{\sigma}_e^2 = 68{,}519.9$	$\hat{\sigma}_s^2 = 3{,}822.3$	$\hat{h}^2 = 0.211 \pm 0.018$
Milk fat:			
Between sires	812	18,203.3	$\sigma_e^2 + 24.62\ \sigma_s^2$
Within sires	19,266	8,891.0	σ_e^2
	$\hat{\sigma}_e^2 = 8{,}891.0$	$\hat{\sigma}_s^2 = 378.2$	$\hat{h}^2 = 0.163 \pm 0.016$

smaller environmental variance in the population.

When sires with 2 and 3 daughters in 2 herds at least, were included in the analyses here, the heritability estimates were found to be slightly higher, 0.284, 0.240, 0.232 and 0.188, respectively (Table 5). The two sets of estimates are summarized in Table 3.

Table 3. Estimates of heritability

Trait	Heritability[1]	Heritability[2]
Mature equivalent records:		
Milk	0.257 ± 0.019	0.284 ± 0.020
Milk fat	0.200 ± 0.017	0.240 ± 0.020
Deviation records:		
Milk	0.211 ± 0.018	0.232 ± 0.020
Milk fat	0.163 ± 0.016	0.188 ± 0.016

[1]Sires with at least four daughters with no restriction on number of herds they are in.

[2]Sires with at least two daughters in at least two herds.

These results disagree with Butcher's (1965) findings. In his data, the values of h^2 decreased when sires with daughters in less than four herds were included in the analyses. All milk and milk fat production was coded to nearest 10 lbs. Therefore all the mean squares need to be multiplied by 100. Thus, the total variance for deviation milk is, (Table 2), (68,519.9 + 3,822.3) X 100 i.e. about 7,200,000 which is considerably larger than 5 to 6 million - the figure generally found in literature. Similarly for milk fat deviation the magnitude of the

variance is higher, approximately 920,000 vs. 550,000 to 700,000 (Thomson 1968, Butcher and Freeman 1968, Ramsay 1966, Blanchard 1965, etc.). The sire component of variance is larger, but not in the same proportion as compared to the findings of the above workers. This is why heritability estimates are smaller than those generally found in literature.

B. Environmental Correlations

1. The difference in intra-class correlations among daughters in one herd versus in many herds

The problem of accounting for environmental correlation, c^2, among paternal half-sibs arises when several half-sibs are in the same herd. The value of c^2 is maximum when all daughters of a sire are in one herd, and is zero when each daughter is in a different herd. Heritability estimates were obtained for these two situations leading to an estimate of c^2. The situations were:

a. When all daughters of a sire were in one herd. This was achieved by selecting at random, one herd per sire. All the daughters in the selected herd were used in the analyses of variance.

b. When each daughter of a sire was in a different herd. This was achieved by selecting at random, one daughter per herd when there was more than one.

The results of between and within sire analyses of variance for records of paternal half-sibs selected in the two situations, a) and b), are given in Tables 4 and 5, respectively. The difference between the two

estimates of heritability is $4c^2$.

While deviation records are designed to eliminate environmental effects related to herd and year-season of freshening, estimates based on mature-equivalent records still contain any herd, year or season effects. The herd effect for different traits were obtained by computing separate analyses of variance on the first lactation records of the daughters of sires that were used in single herds only. The results of these analyses are present in Table 6.

The intra-class correlation for mature-equivalent records obtained from Table 4 was adjusted by subtracting from it the intra-class correlation among herds obtained from the analyses of variance for single herd sires (Table 6). Thus, the adjusted intra-class correlation for mature-equivalent milk was:

$$\frac{30806.4}{(30806.4 + 61626.0)} - \frac{21464.2}{(21464.2 + 11211.4 + 53602.7)} = 0.0845$$

Four times this correlation estimates heritability as given in Table 7. Similar adjustments were made for deviation records. Results obtained later (Tables 9, 10 and 11) indicated that deviating records does not remove the herd and year-season effects completely. When each daughter is in a different herd such adjustment is not required. The estimates of c^2 obtained in this manner are given in Table 7, where $\hat{h}^2_a$ is the heritability estimated under situation a) and $\hat{h}^2_b$ is the heritability estimated under situation b). The estimates of c^2 seem small.

Table 4. Analysis of variance - heritability and correlation among half-sibs in the same herd

MATURE-EQUIVALENT RECORDS

Source of variation	d.f.	M.S.	E.M.S.
Milk:			
Between sires	1,559	278,601.5	$\sigma_e^2 + 7.04\ \sigma_s^2$
Within sires	9,438	61,626.0	σ_e^2
	$\hat{\sigma}_e^2 = 61,626.0$	$\hat{\sigma}_s^2 = 30,806.4$	$\hat{t} = 0.333 \pm 0.011$
Milk fat:			
Between sires	1,559	39,922.1	$\sigma_e^2 + 7.04\ \sigma_s^2$
Within sires	9,438	7,754.9	σ_e^2
	$\hat{\sigma}_e^2 = 7,754.9$	$\hat{\sigma}_s^2 = 4,567.1$	$\hat{t} = 0.371 \pm 0.011$

DEVIATION RECORDS

Source of variation	d.f.	M.S.	E.M.S.
Milk:			
Between sires	1,559	104,468.4	$\sigma_e^2 + 7.04\ \sigma_s^2$
Within sires	9,438	62,992.8	σ_e^2
	$\hat{\sigma}_e^2 = 62,992.8$	$\hat{\sigma}_s^2 = 5,888.7$	$\hat{t} = 0.086 \pm 0.007$
Milk fat:			
Between sires	1,559	13,625.2	$\sigma_e^2 + 7.04\ \sigma_s^2$
Within sires	9,438	7,954.9	σ_e^2
	$\hat{\sigma}_e^2 = 7,954.9$	$\hat{\sigma}_s^2 = 805.1$	$\hat{t} = 0.092 \pm 0.008$

Table 5. Analysis of variance - heritability and correlation among half-sibs in different herds

MATURE-EQUIVALENT RECORDS

Source of variation	d.f.	M.S.	E.M.S.
Milk:			
Between sires	1,559	156,303.9	$\sigma_e^2 + 11.54\ \sigma_s^2$
Within sires	20,745	82,848.8	σ_e^2

$\hat{\sigma}_e^2 = 82,848.8$ $\hat{\sigma}_s^2 = 6,367.6$ $\hat{t} = 0.071 \pm 0.005$ $\hat{h}^2 = 0.284 \pm 0.020$

Milk fat:			
Between sires	1,559	18,976.7	$\sigma_e^2 + 11.54\ \sigma_s^2$
Within sires	20,745	10,956.4	σ_e^2

$\hat{\sigma}_e^2 = 10,956.4$ $\hat{\sigma}_s^2 = 695.3$ $\hat{t} = 0.060 \pm 0.005$ $\hat{h}^2 = 0.240 \pm 0.020$

DEVIATION RECORDS

Source of variation	d.f.	M.S.	E.M.S.
Milk:			
Between sires	1,559	118,392.8	$\sigma_e^2 + 11.54\ \sigma_s^2$
Within sires	20,745	69,326.0	σ_e^2

$\hat{\sigma}_e^2 = 69,326.0$ $\hat{\sigma}_s^2 = 4,253.5$ $\hat{t} = 0.058 \pm 0.005$ $\hat{h}^2 = 0.232 \pm 0.020$

Milk fat:			
Between sires	1,559	14,175.7	$\sigma_e^2 + 11.54\ \sigma_s^2$
Within sires	20,745	9,003.9	σ_e^2

$\hat{\sigma}_e^2 = 9,003.9$ $\hat{\sigma}_s^2 = 448.3$ $\hat{t} = 0.047 \pm 0.004$ $\hat{h}^2 = 0.188 \pm 0.016$

Table 6. Analysis of variance - single herd sires only

Source of variation	d.f.	M.S.	E.M.S.
Mature-equivalent milk:			
Between herds	730	656,685.7	$\sigma_e^2 + 11.91\ \sigma_s^2 + 1.87\ \sigma_H^2$
Between sires/herds	691	172,083.2	$\sigma_e^2 + 10.57\ \sigma_s^2$
Within sires/herds	14,594	53,602.7	σ_e^2
	$\hat{\sigma}_e^2 = 53,602.7$	$\hat{\sigma}_s^2 = 11,211.4$	$\hat{\sigma}_H^2 = 21,462.2$
Mature-equivalent milk fat:			
Between herds	730	89,159.2	$\sigma_e^2 + 11.91\ \sigma_s^2 + 1.87\ \sigma_H^2$
Between sires/herds	691	21,989.6	$\sigma_e^2 + 10.57\ \sigma_s^2$
Within sires/herds	14,594	6,843.0	σ_e^2
	$\hat{\sigma}_e^2 = 6,843.0$	$\hat{\sigma}_s^2 = 1,433.3$	$\hat{\sigma}_H^2 = 2,982.5$
Deviation milk:			
Between herds	730	122,845.9	$\sigma_e^2 + 11.91\ \sigma_s^2 + 1.87\ \sigma_H^2$
Between sires/herds	691	101,494.8	$\sigma_e^2 + 10.57\ \sigma_s^2$
Within sires/herds	14,594	54,889.3	σ_e^2
	$\hat{\sigma}_e^2 = 54,889.3$	$\hat{\sigma}_s^2 = 4,410.1$	$\hat{\sigma}_H^2 = 704.9$
Deviation milk fat:			
Between herds	730	15,018.7	$\sigma_e^2 + 11.91\ \sigma_s^2 + 1.87\ \sigma_H^2$
Between sires/herds	691	12,831.0	$\sigma_e^2 + 10.57\ \sigma_s^2$
Within sires/herds	14,594	7,034.0	σ_e^2
	$\hat{\sigma}_e^2 = 7,034.0$	$\hat{\sigma}_s^2 = 548.6$	$\hat{\sigma}_H^2 = 66.3$

Table 7. Estimation of environmental correlation among half-sibs from heritability estimates obtained under situations a) and b)

Trait	$\hat{h}^2_a$	$\hat{h}^2_b$	$4\hat{c}^2$	$\hat{c}$
Mature-equivalent records:				
Milk	0.338	.284	.054	.014
Milk fat	0.423	.240	.183	.046
Deviation records:				
Milk	0.295	.232	.063	.016
Milk fat	0.333	.188	.145	.036

Various estimates of intra-class correlations obtained under situations a) and b) are summarized in Table 8. The figures in parentheses are by Thomson (1968) and those in brackets are by Van Vleck (1966b). Whereas the correlation estimated for mature-equivalent milk, when half-sibs are in the same herd, $\hat{t}$ = 0.333, is similar to the values 0.335 obtained by Thomson (1968), the estimate obtained from the analysis when half-sibs were in different herds is slightly lower (0.071 vs. 0.104). The corresponding values for deviation milk records estimated here, are considerably lower than those of Thomson, (0.086 and 0.058 vs. 0.176 and 0.074).

Van Vleck (1966b) computed similar correlations among the daughter pairs in the same and different herds to be 0.168 and 0.086, respectively. Those are for deviation milk and are considerably higher than those obtained here, 0.074 and 0.058 (1/4 h^2). Consequently, his estimate of c^2 also is of larger magnitude, 0.082 vs. 0.016.

Table 8. Estimates of intra-class correlations obtained under situations a) and b)

Trait	a)	b)
Mature-equivalent records:		
Milk	0.333 (0.335)	0.071 (0.104)
Milk fat	0.371	0.060
Deviation records:		
Milk	0.086 (0.176) [0.168]	0.058 (0.074) [0.086]
Milk fat	0.092	0.047

2. Intra-class correlation due to herds, year-seasons and herd-year-seasons

A lactation record of a cow was represented by the following model:

$$Y_{ijkl} = \mu + a_i + h_j + s_k + e_{ijkl}$$

where,

Y_{ijkl} is the first lactation record of the l^{th} daughter, of the k^{th} sire, made in the j^{th} herd and started in the i^{th} year-season,

μ is the population mean,

a_i is an effect common to all records in the i^{th} year-season,

h_j is an effect common to all records made in the j^{th} herd,

s_k is an effect common to all records made by daughters of the k^{th} sire, and

e_{ijkl} is an effect peculiar to the record of the l^{th} daughter, of the

k^{th} sire, made in the j^{th} herd and started in the i^{th} year-season ($l = 1, 2, \ldots, n_{jk}$).

All effects, except the mean, were assumed to be random, independent and normally distributed with means equal to zero and variances, σ_a^2, σ_h^2, σ_s^2 and σ_e^2, respectively. The covariance $\sigma_{e_{ijkl'}e_{ij'k'l}}$ was assumed to be equal to zero for all k not equal to k' and/or j not equal to j'. In the case of $k = k'$ and $j = j'$, the covariance $\sigma_{e_{ijkl'}e_{ijkl}}$ was assumed to be equal to $\rho^2\ \sigma_e^2$, where ρ^2 in the additional environmental correlation between half-sibs calving in the same herd.

If only first lactation records are considered, then the average of all the daughters of a particular sire, say k', will be:

$$\bar{Y}_{..k'.} = \frac{1}{n_{..k'}} \sum\sum\sum_{ijl} Y_{ijk'l} = \mu + \frac{\sum_i a_i}{n_{..k'}} + \frac{\sum_j h_j}{n_{..k'}} + s_{k'} + \frac{\sum\sum\sum_{ijl} e_{ijk'l}}{n_{..k'}},$$

which has an expected value of:

$$E[\bar{Y}_{..k'.}] = \mu + s_{k'},$$

and a variance (Thomson, 1968) of:

$$V[\bar{Y}_{..k'.}] = V[Y_{ijkl}] \frac{4 + (n_{..k'}-1)\ h^2 + 4 \sum\sum_{ij} \left(\frac{n_{ijk'}(n_{ijk'}-1)}{n_{..k'}}\right)(s+\rho^2)}{4\ n_{..k'}} +$$

$$\frac{4 \sum_j \left(\sum_{i \neq i'} \frac{(n_{ijk'})(n_{i'jk'})}{n_{..k'}}\right)(t+\rho^2) + 4 \sum_i \left(\sum_{j \neq j'} \frac{(n_{ijk'})(n_{ij'k'})}{n_{..k'}}\right)(w)}{4\ n_{..k'}}$$

where,

$$h^2 = 4\,\sigma_s^2 / (\sigma_a^2 + \sigma_h^2 + \sigma_s^2 + \sigma_e^2),$$

$$s = (\sigma_a^2 + \sigma_h^2) / (\sigma_a^2 + \sigma_h^2 + \sigma_s^2 + \sigma_e^2),$$

$$t = \sigma_h^2 / (\sigma_a^2 + \sigma_h^2 + \sigma_s^2 + \sigma_e^2), \text{ and}$$

$$w = \sigma_a^2 / (\sigma_a^2 + \sigma_h^2 + \sigma_s^2 + \sigma_e^2).$$

If all the half-sibs are in different herds and different year-seasons, the above equation reduces to:

$$V[\bar{Y}_{..k'.}] = V[Y_{ijkl}] \left[\frac{4 + (n_{..k'} - 1)\, h^2}{4\, n_{..k'}}\right].$$

The records of any two half-sisters must always fall in one of the four classifications given in the following table:

Table 9. Components of covariance among paternal half-sibs in the same and different herds and year-seasons

Classification	Component of covariance
Same herd - same year-season	$\sigma_s^2 + \sigma_a^2 + \sigma_h^2 + \rho^2\ \sigma_e^2$
Same herd - different year-seasons	$\sigma_s^2 + \sigma_h^2 + \rho^2\ \sigma_e^2$
Different herds - same year-season	$\sigma_s^2 + \sigma_a^2$
Different herds - different year-seasons	σ_s^2

In addition to the sire component, σ_s^2, half-sibs in the same year-season have a year-season component, σ_a^2, in common. The half-sibs in the same herd have a herd component, σ_h^2 included in the covariance. In

addition, a term $\rho^2 \sigma_e^2$ is included to represent the additional covariance among half-sibs being in the same herd. This term would not appear in the covariance between half-sibs in different herds whether or not they calve in the same year-season.

The environmental correlations, s, t and w, were estimated as the intra-class correlations derived from a between and within herd-year-season, herd and year-season analyses of variance for both mature-equivalent and deviation records. All the records were included in the analyses for the correlation (s) among daughters in the same herd and same year-season. However, in the analyses for the correlation (t) among daughters in the same herd but different year-seasons, and the correlation (w) among daughters in the same year-season but different herds, one daughter's record was selected at random from each herd-year-season. The herd component in a between and within herds analyses then will be expected to contain only the differences between herds. Similarly, the year-season component of variance in a between and within year-season analyses will be expected to contain only the differences between year-seasons. Results of these analyses are given in Tables 10, 11 and 12.

The analyses of mature-equivalent records of milk production gave estimates of s that were slightly higher than those reported by Thomson (1968) and Bereskin (1963). Thomson estimated the correlation among records of cows freshening in a year-season at 0.328. He used 158,236 records from data obtained from the Holstein-Friesian Association of America. These records were made by 62,389 cows scattered all over the country during the years 1952 through 1961. This value is only slightly lower than the correlation of 0.357 obtained in this study. The

correlation for milk fat found here was even higher, 0.381. Thomson (1968) used five months rolling seasons and a season was defined as one month for the purpose of the analyses. In the present study fixed seasons were used; those were - May through September and October through the following April. Bereskin (1963) estimated the correlation to be 0.332 which is close to that of Thomson (1968). He used 33,139 Holsteins records obtained from the Iowa DHIA Central Processing Center. Fixed seasons, May through September and October through the following April, were used in his analyses also.

The results of the analyses of deviation records indicated, as often suspected, that deviating records did not remove herd- and year-season-effects completely, at least in these data. All the three correlations, s, t and w, are above zero for both milk and milk fat production (milk: 0.045, 0.038 and 0.015; milk fat: 0.060, 0.043 and 0.010).

Thomson (1968) estimated s and t for deviation milk records to be essentially zero. His estimate of w, the correlation among cows in the same year-season but different herds, was above zero but only slightly so (0.003). He concluded that deviating records was an effective method of removing herd- and year-season-effects. The discrepancies between the results obtained in this study and those reported by Thomson (1968) probably can be attributed to the real differences between the two sets of data. His population was scattered all over the United States and included a lower proportion of sires used in many herds than these data from the eight states.

It is possible that the paternal half-sisters that calve in the same herd are treated differently than the other cows in the herd. This could

Table 10. Analysis of variance - correlation (s) among cows in the same herd and year-season

MATURE-EQUIVALENT RECORDS

Source of variation	d.f.	M.S.	E.M.S.
Milk:			
Between herd-year-season	14,491	180,113.9	$\sigma_e^2 + 3.81\ \sigma_{hys}^2$
Within herd-year-season	40,678	57,860.1	σ_e^2

$\hat{\sigma}_e^2 = 57,860.1$ $\hat{\sigma}_{hys}^2 = 32,115.4$ $\hat{s} = 0.357 \pm 0.005$

Milk fat:			
Between herd-year-season	14,491	24,359.2	$\sigma_e^2 + 3.81\ \sigma_{hys}^2$
Within herd-year-season	40,678	7,260.0	σ_e^2

$\hat{\sigma}_e^2 = 7,260.0$ $\hat{\sigma}_{hys}^2 = 4,491.9$ $\hat{s} = 0.381 \pm 0.005$

DEVIATION RECORDS

Source of variation	d.f.	M.S.	E.M.S.
Milk:			
Between herd-year-season	14,491	75,094.7	$\sigma_e^2 + 3.81\ \sigma_{hys}^2$
Within herd-year-season	40,678	63,592.1	σ_e^2

$\hat{\sigma}_e^2 = 63,592.1$ $\hat{\sigma}_{hys}^2 = 3,021.7$ $\hat{s} = 0.045 \pm 0.004$

Milk fat:			
Between herd year-season	14,491	9,925.1	$\sigma_e^2 + 3.81\ \sigma_{hys}^2$
Within herd-year-season	40,678	7,983.3	σ_e^2

$\hat{\sigma}_e^2 = 7,983.3$ $\hat{\sigma}_{hys}^2 = 510.1$ $\hat{s} = 0.060 \pm 0.004$

Table 11. Analysis of variance - correlation (t) among cows in the same herd but different year-seasons

MATURE-EQUIVALENT RECORDS

Source of variation	d.f.	M.S.	E.M.S.
Milk:			
Between herds	2,336	181,493.0	$\sigma_e^2 + 6.23\ \sigma_h^2$
Within herds	12,165	71,425.5	σ_e^2

$\hat{\sigma}_e^2 = 71{,}425.5$ $\hat{\sigma}_h^2 = 17{,}677.8$ $\hat{t} = 0.198 \pm 0.009$

Milk fat:			
Between herds	2,336	24,756.3	$\sigma_e^2 + 6.23\ \sigma_h^2$
Within herds	12,165	9,123.0	σ_e^2

$\hat{\sigma}_e^2 = 9{,}123.0$ $\hat{\sigma}_h^2 = 2{,}510.8$ $\hat{t} = 0.216 \pm 0.008$

DEVIATION RECORDS

Source of variation	d.f.	M.S.	E.M.S.
Milk:			
Between herds	2,336	81,277.6	$\sigma_e^2 + 6.23\ \sigma_h^2$
Within herds	12,165	65,138.3	σ_e^2

$\hat{\sigma}_e^2 = 65{,}138.3$ $\hat{\sigma}_h^2 = 2{,}592.1$ $\hat{t} = 0.038 \pm 0.006$

Milk fat:			
Between herds	2,336	10,629.0	$\sigma_e^2 + 6.23\ \sigma_h^2$
Within herds	12,165	8,295.5	σ_e^2

$\hat{\sigma}_e^2 = 8{,}295.5$ $\hat{\sigma}_h^2 = 374.8$ $\hat{t} = 0.043 \pm 0.006$

Table 12. Analysis of variance - correlation (w) among cows in the same year-season but in different herds

MATURE-EQUIVALENT RECORDS

Source of variation	d.f.	M.S.	E.M.S.
Milk:			
Between year-seasons	22	1,572,581.7	$\sigma_e^2 + 751.10\ \sigma_{ys}^2$
Within year-seasons	17,884	87,817.3	σ_e^2
	$\hat{\sigma}_e^2 = 87,817.3$	$\hat{\sigma}_{ys}^2 = 1,976.8$	$\hat{w} = 0.022 \pm 0.007$
Milk fat:			
Between year-seasons	22	167,441.0	$\sigma_e^2 + 751.10\ \sigma_{ys}^2$
Within year-seasons	17,884	11,563.2	σ_e^2
	$\hat{\sigma}_e^2 = 11,563.2$	$\hat{\sigma}_{ys}^2 = 207.5$	$\hat{w} = 0.018 \pm 0.006$

DEVIATION RECORDS

Source of variation	d.f.	M.S.	E.M.S.
Milk:			
Between year-seasons	22	829,393.8	$\sigma_e^2 + 751.10\ \sigma_{ys}^2$
Within year-seasons	17,884	68,106.4	σ_e^2
	$\hat{\sigma}_e^2 = 68,106.4$	$\hat{\sigma}_h^2 = 1,013.6$	$\hat{w} = 0.015 \pm 0.005$
Milk fat:			
Between year-seasons	22	76,366.0	$\sigma_e^2 + 751.10\ \sigma_{ys}^2$
Within year-seasons	17,884	8,699.0	σ_e^2
	$\hat{\sigma}_e^2 = 8,699.0$	$\hat{\sigma}_h^2 = 90.1$	$\hat{w} = 0.010 \pm 0.003$

occur if they were all housed together in the same area of the barn, or given a little extra care. Thus, in addition to being half-sibs and being in the same herd, they could have an extra likeness (correlation). The correlation among these half-sibs would, therefore, be higher than expected otherwise, due to a correlated environmental effect. The correlation between two half-sisters in different herds is expected to be 1/4 heritability, while the correlation between two half-sisters in the same herd also includes t, a correlation due to being in the same herd, and an additional correlation ρ^2, due to environmental conditions common to the pair. In other words:

1. Different herds, correlation = 1/4 h^2, and
2. Same herds, correlation = 1/4 h^2 + t + ρ^2.

The results of the analyses in Tables 4, 5, 6 and 11 were used to evaluate these additional correlations and are summarized in Table 13.

Table 13. The additional correlations (ρ^2) among paternal half-sisters in the same herd

Trait	ρ^2(1)	ρ^2(2)
Mature-equivalent records:		
Milk	0.064	0.059
Milk fat	0.095	0.068
Deviation records:		
Milk	-0.010	0.016
Milk fat	0.002	0.024

The environmental correlation between half-sibs in the same herd for mature-equivalent milk is 0.333 - 0.071 = 0.262. The correlation among half-sibs in the same herd, t, is 0.198. Thus, the difference, or 0.064, is the additional correlation among half-sibs due to common intra-herd environmental effects. Similar estimates for other traits were obtained and are present in the first column, ρ^2 (1), of Table 13.

The same additional correlations were estimated using a slightly different procedure. The intra-sire correlation (sire within herd) estimated in the single herd sire analyses of variance (Table 6), i.e. $\hat{\sigma}^2_s / (\hat{\sigma}^2_s + \hat{\sigma}^2_h + \hat{\sigma}^2_e)$, contains 1/4 heritability and ρ^2. Since the herd effect has been removed in the analysis, no t appears in the correlation among these half-sibs. If the correlation among the half-sibs in different herds, i.e. $\hat{\sigma}^2_s / (\hat{\sigma}^2_s + \hat{\sigma}^2_e)$, which is expected to be 1/4 heritability, obtained from Table 5 is subtracted from the above correlation, the remainder must estimate ρ^2. Thus, for mature-equivalent milk the additional correlation is 0.130 - 0.071 = 0.059. The denominators in the two intra-class correlations differ in terms of components of variance because of the models used. In the case of intra-sire within herd correlation, σ^2_h is included in the σ^2_e, although there is always a small fraction of it confounded in the sire component (σ^2_s). So the denominators contain the total variance in both cases.

Similar estimates for other traits were obtained and are given in the second column, ρ^2 (2), of Table 13.

The basic difference between the two estimates of ρ^2 is due to the manner in which they were estimated. For estimating ρ^2 (1), the same sires were represented in the analyses of half-sibs in the same herd and in the

analyses of half-sibs in different herds. However, for estimating $\rho^2(2)$, the sires were different in the two sets of analyses. Only single herd sires' daughters' records were used to get the correlation among half-sibs in the same herd.

The estimates for the additional correlation obtained in this study are different from those reported by Thomson (1968). His estimate for mature-equivalent milk (0.005) is considerably lower; however, for deviation milk he obtained a very large correlation (0.102). He concluded that this large difference could be due, at least in part, to the fact that the correlation among half-sibs was estimated from first lactation records only, while the correlation among cows in the same herd was estimated from all records. Due to selection, the correlation among cows in the same herd would be expected to be higher when all records are considered rather than if only first lactations are considered. This should make the estimate of ρ^2 for mature-equivalent records very low. On the other hand in the case of deviation records the herdmate average contain a larger fraction of older cows, because paternal half-sibs are omitted from the herdmate average. Any errors in age correction factors used would tend to inflate the sire component more for the analyses with all records in one herd than the analyses with each record in different herds. The result would be over estimation of ρ^2 for deviation record but not for mature-equivalent records.

In the light of the above conclusions the estimates obtained in the present study were as expected. Since only first lactation records were used in all analyses, no over-estimation in the mature-equivalent records analyses or under-estimation in deviation records analyses could occur.

Estimates of ρ^2 for deviation records were expected to be lower as compared to those for mature-equivalent records and they were here.

3. Within herd and sire intra-class correlation

The intra-herd correlation (within a sire) is an estimate of the environmental correlation among daughters in the same herd. Since the herd and year-season effects are assumed to be removed but are not always perfectly, by deviating records, the intra-herd correlation among paternal half-sibs in the same herd should estimate the total environmental correlation among half-sibs plus any such correlation remaining because of inadequate correction of data. The following model was assumed to describe the individual records of progeny of a sire:

$$Y_{ij} = \mu + h_i + e_{ij}$$

where,

Y_{ij} is the first lactation record of the j^{th} daughter in the i^{th} herd,

μ is the population mean,

h_i is an effect common to all records made in the i^{th} herd, and

e_{ij} is an effect peculiar to the record of the j^{th} daughter in the i^{th} herd ($j = 1, 2, \ldots, n_i$).

From a between and within herds analysis of variance, the intra-herd correlation among the daughters of a sire can be computed as $\sigma_h^2 / (\sigma_h^2 + \sigma_e^2)$. In order to pool all the sires together, the above model was expanded to:

$$Y_{ijk} = \mu + s_i + h_{j(i)} + e_{ijk}$$

where,

Y_{ijk} is the first lactation record of the k^{th} daughter, in the j^{th} herd, nested in the i^{th} sire,

μ is the population mean,

s_i is an effect of the i^{th} sire,

$h_{j(i)}$ is the effect of the j^{th} herd nested in the i^{th} sire, and

e_{ijk} is a random effect peculiar to the first record of the k^{th} daughter, in the j^{th} herd, nested in the i^{th} sire ($k = 1, 2, \ldots, n_{ij}$).

All effects, except the mean, were assumed to be random independent and normally distributed with means equal to zero and variances σ_s^2, $\sigma_{h(s)}^2$ and σ_e^2, respectively.

The estimates of the sire-by-herd interaction, reported in the literature, have been small and assumed to be of little practical importance, indicating that the sires are ranked essentially the same in different herds. (Hickman and Henderson, 1955; Legates, et al., 1956; Mason and Robertson, 1956; Wadell and McGilliard, 1959; Robertson, et al., 1960; Van Vleck, et al., 1961; Burdick and McGilliard, 1963; Van Vleck, 1963; and Kelleher, 1964.) Since such is the case in general, the sire-by-herd interaction effects were assumed to be unimportant and, therefore, the interaction term was not included in the above model.

The environmental correlation c^2 was estimated as the intra-herd (nested in sires) correlation derived from a between herds, between herds nested in sires and within herds nested in sires analyses of variance for

deviation milk and milk fat records.

An obvious expectation is that the artificial insemination service (AI) sires, since they have daughters scattered over several herds, have a smaller environmental correlation among daughters in the same herd as compared to natural service (NS) sires which, in general, have daughters that produce only in a few herds. To study this difference in environmental correlations, if there was any, two sets of data were obtained: Set 1 included first lactation records of the daughters of sires that had 50 or more daughters scattered over several herds, that is AI sire's daughters, and Set 2 included first lactation records of the daughters of sires that had daughters only in 2 to 6 herds, that is NS sire's daughters. There were 197 sires with 34,753 daughters in Set 1 and 829 sires with 12,774 daughters in Set 2. Separate analyses of variance for the two sets are given in Table 14.

The values of c^2 for deviation milk and milk fat, were computed to be 0.103 and 0.113, respectively among AI sired daughters, and 0.112 and 0.120, respectively among NS sired daughters. These estimates of c^2 and the various components of variance are summarized in Table 15.

Table 14. Analyses of variance - intra-herd correlation among daughters of AI sires and NS sires (deviation records)

Source of variation	d.f.	M.S.		E.M.S.
		Milk	Milk fat	
AI sires:				
Between sires	196	630,730.5	66,613.5	$\sigma_e^2 + 7.04\ \sigma_{h(s)}^2 + 175.55\ \sigma_s^2$
Between herds within sires	14,208	75,646.4	9,886.9	$\sigma_e^2 + 2.35\ \sigma_{h(s)}^2$
Within herds within sires	20,348	58,936.0	7,545.6	σ_e^2
NS sires:				
Between sires	828	139,498.4	19,262.7	$\sigma_e^2 + 10.36\ \sigma_{h(s)}^2 + 15.4\ \sigma_s^2$
Between herds within sires	1,458	84,376.6	10,919.7	$\sigma_e^2 + 2.87\ \sigma_{h(s)}^2$
Within herds within sires	10,488	61,948.0	7,855.7	σ_e^2

Table 15. Estimates of variance components and the intra-herd correlation (deviation records)

Variance and correlation	Components			
	AI sires		NS sires	
	Milk	Milk fat	Milk	Milk fat
$\hat{\sigma}^2_s$	2,971.9	296.5	-232.4	21.0
$\hat{\sigma}^2_{h(s)}$	7,115.5	997.0	7,827.9	1,069.4
$\hat{\sigma}^2_e$	58,936.0	7,545.6	61,948.0	7,855.7
$\hat{c}^2$	0.103 $\pm$ 0.007	0.113 $\pm$ 0.007	0.112 $\pm$ 0.015	0.120 $\pm$ 0.015

While estimate for deviation milk is a little higher than that reported by Van Vleck (1966b), 0.082, it is slightly lower than the value, 0.140, currently being used in USDA sire evaluation method (Plowman and McDaniel, 1968). Some of the data used in the present study may have been included in the USDA estimate of c^2, but much of the data and the method of estimation were different. This method of computing c^2 is rather simple and seems directly applicable to the current USDA sire evaluation program. It is the within a herd variation caused by favorable or unfavorable treatment given to daughters of some particular sire(s) and the failure of statistical corrections for environmental effects to be perfect, that give rise to c^2 (actually this is the kind of environmental correlation that has practical importance), rather than the mere fact that the daughters are in the same herd.

The estimate of c^2 reported by Van Vleck (1966b) and Plowman and McDaniel (1968) were computed by using records of all daughters of AI and

NS sires combined. In the same study, Van Vleck estimated c^2 values separately for AI sired pairs and NS sired pairs in the same herd. He found no evidence to indicate that the environmental correlation was lower for AI sired pairs than the NS sired pairs, rather observed the reverse, 0.130 vs. 0.088. In the present study the estimates of c^2 obtained for AI sired daughters in the same herd are lower than those estimated for NS sired daughters in the same herd, as expected, but by only a very small amount, 0.103 vs. 0.112 for deviation milk and 0.113 vs. 0.120 for deviation milk fat.

However, it must be pointed out that the correlation s, t and w for deviated records, as computed previously (Tables 10, 11 and 12), were above zero for the data used here, indicating that deviation records would not remove herd- and year-season-effects completely. Therefore, the true environmental correlations among half-sibs in the same herd, computed by the present procedure probably are over-estimated. But, they should be over-estimated by only a small fraction because the correlations s, t and w were very small.

4. Simultaneous estimation of c^2 and h^2

A modified Gauss-Newton method for the fitting of a non-linear regression function by least squares has been developed by Hartley (1961). This method, which is an iterative procedure to estimate parameters from a non-linear model, was used to compute least squares estimates of c^2 and h^2. A sketch of the theoretical considerations in the procedure follows:

Given n sets of observed (k + 1) - tuples y_h; x_{1h}, x_{2h}, - - - x_{kh} (h = 1, 2, . . ., n), such that x's are independent variables, y is

dependent, and:

$$y = f\ (x;\ \theta) = f\ (x_1, \ldots, x_k;\ \theta_1, \theta_2, \ldots, \theta_m); \qquad 16$$

the classical least squares problem is to determine a set of θ_i the unknown parameters, for which the sum of squares:

$$Q\ (\theta) = \sum_{h=1}^{n} [y_h - f\ (x_h;\ \theta)]^2 = \text{Min.} \qquad 17$$

The function 16 is expanded in a 1^{st} order Taylor series:

$$f\ (x_h;\ \theta) = f\ (x_h;\ \theta_o) + \sum_{i=1}^{m} f_i\ (x_h;\ \theta_o)\ (\theta_i - \theta_{io}) \qquad 18$$

where $(\theta_i - \theta_{io})$ is the familiar Newton coefficient D_i. Thus, $Q\ (\theta)$ can now be expressed as a linear model with respect to the unknown Newton coefficients. These coefficients may converge to zero if the starting values are correct, that is:

$$\lim_{k \to \infty} \left| \theta_i^{k+1} - \theta_i^k \right| = 0.$$

Thus:

$$Q\ (\theta) = \sum_h [(y_h - f\ (x_h;\ \theta_o)) - \sum_i f_i\ (x_h;\ \theta_o)\ D_i]^2 \qquad 19$$

yielding the least squares equations:

$$2 \sum_{j=1}^{m} [\sum_h f_i\ (x_h;\ \theta_o)\ f_j\ (x_h;\ \theta_o)]\ D_j = -\ Q_i\ (x;\ \theta_o) \qquad 20$$

where $Q_i\ (x;\ \theta_o)$ is given by the first derivative of 17 with regard to θ_i. Then the j^{th} set of equations is:

$$\sum_{j=1}^{m} \left[\sum_{h} f_i (x_h; \theta_o) f_j (x_h; \theta_o)\right] D_j =$$

$$\sum_{h=1} (y_h - f (x_h; \theta_o)) f_j (x_h; \theta_o). \qquad 21$$

By assumption the determinant of the linear equations 21 has rank m and thus can always be solved, yielding the element D_i of the vector D as solutions.

It is at this point that Hartley (1961) modified the usual 'Gauss-Newton' method of iterative solution. Consider the function:

$$Q (v) = Q (x, \theta + v D), \text{ for } 0 \leq v \leq 1, \qquad 22$$

and denote by v' the value of v for which Q (v) is minimum on the interval $0 \leq v \leq 1$. Defining the vector $\theta_1 = \theta_o + v' D$ with elements $\theta_{io} + v' D_i$, obviously:

$$Q (x, \theta_1) \leq Q (x, \theta_o) < Q; \qquad 23$$

so that θ_1 clearly lies in the interior of bounded convex set (S). This computation is now repeated at θ_1 and so on resulting in a sequence of vectors θ_t, t = 1, 2, . . ., all within the bounded convex set.

In other words using D_i iterate:

$$\theta_i^{(k+1)} = \theta_i^{(k)} + v D_i$$

until $\left| \theta_i^{(k+1)} - \theta_i^{(k)} \right| < \delta$ where δ is a pre-established tolerance.

The fraction v according to Hartley insures convergence and probably hastens it. For practically all problems of non-linear regression, the original sequence of θ's will converge to the solution θ^*. It is highly

improbable that there will be a regression surface and a set of observed x_h and y_h such that Q has two stationary points yielding precisely the same value of Q.

This iterative process converges to the vector θ^{**} yielding the absolute minimum of Q i.e. the unique solution, provided that the starting vector θ_o is in the 'neighborhood region' of the bounded convex that contains the absolute minimum solution.

Atkinson (1966), using an interpretation of Hartley's modified method, designed a Fortran IV program to estimate θ's from relatively simple non-linear models. If $\theta^{(k)}$ is defined as being a function of v, then Q (θ) is also a function of v; that is,

$$Q(v) = A v^2 + B v + C \qquad 24$$

The problem is to find the minimum of the quadratic in v fitted through the three points v = 0, v = 1/2 and v = 1. Since the minimum of a quadratic is given by -B / 2A, solving for A and B above:

$$C = Q(0) = Q(\theta^k)$$

$$A/4 + B/2 + C = Q(1/2) = Q(\theta^k + 1/2\ D)$$

$$A + B + C = Q(1) = Q(\theta^k + D)$$

yielding:

$$A = 2\,[Q\,(1) - 2\,Q\,(1/2) + Q\,(0)]$$

$$B = 4\,Q\,(1/2) - Q\,(1) - 3\,Q\,(0)$$

$$C = Q\,(0)$$

Substituting these values into the min. of quadratic - B / 2 A gives the minimum of v:

$$v \text{ min.} = 1/2 + 1/4 * (Q\,(0) - Q\,(1)) / (Q\,(1) - 2\,Q\,(1/2) + Q\,(0)).$$

The parabola through Q (0), Q (1/2), Q (1) attains a minimum for the above level of v. This program was used to solve, for c^2 and h^2, the regression formula:

$$b = \frac{M\,h^2}{4 + (M-1)\,h^2 + 4\,\frac{\Sigma m_i (m_i - 1)}{M}\,c^2} \qquad 25$$

On the basis of this regression, the following model was assumed:

$$Y_h = f\,(x_h;\,\theta) = \frac{x_{h1}\,\theta_1}{4 + x_{h2}\,\theta_1 + x_{h3}\,\theta_2} \qquad 26$$

where,

$Y_h = b_h$ is the h^{th} regression of average of daughters in sample S_{Mh} on the average of daughters in sample S_{Nh},

$x_{h1} = M_h$ is the number of daughters in the sample S_{Mh},

$x_{h2} = M - 1$; M is defined as above,

$x_{h3} = 4\,\frac{\Sigma m_i (m_i - 1)}{M}$ is the constant that determines the distribution of daughters acorss the herds; m_i is the number of daughters in the i^{th} herd,

$\theta_1 = h^2$ is the heritability, and

$\theta_2 = c^2$ is the environmental correlation among the daughter (i.e. paternal half-sibs) in the same herd.

In practice, the regression of the breeding value of a sire on the daughter average is usually obtained by using already estimated values of c^2 and h^2 on the right hand side of equation 25. The interest in this study was to estimate c^2 and h^2, therefore, the simple regressions of the average of daughters on the average of an independent sample of daughters, were calculated. Since regressions, b's, are estimated with some unknown and supposedly random error an adjustment for the random error was required. Assume:

$$b_h = \beta_h + \Delta_h$$

where β_h is the true regression and Δ_h the random error. Then the variance of Δ_h and its expectation are:

$$V(\Delta_h) = s^2_{y \cdot x} / \Sigma x^2, \text{ and}$$

$$E[V(\Delta_h)] = \sigma^2 / \Sigma x^2 = \sigma^2 k_h$$

where $s^2_{y \cdot x}$ is the mean square deviation from regression, σ^2 the variance and k_h is the corrected sum of squares of the independent variable used in computation of regression b_h. Incorporating the above adjustment requires minimizing:

$$Q(\theta) = \sum_h \frac{1}{k_h} [b_h - f(x_h; \theta)]^2$$

that is:

$$Q(\theta) = \sum_h \left[\frac{1}{\sqrt{k_h}} b_h - \frac{1}{\sqrt{k_h}} f(x_h; \theta) \right]^2 . \qquad 27$$

Thus, Y_h and X_{h1} in equation 26, are re-defined as:

$$Y_h = b_h / \sqrt{k_h} \qquad 28$$

and

$$X_{h1} = M / \sqrt{k_h} .$$

The daughters of each sire were divided into two groups by dividing at random the number of herds they were used in, into two halves. All the daughters in first half of the herds formed the sample S_M and those in the second half made the sample S_N. When the number of herds was odd, the extra herd was included in S_N. For each sire the averages of deviation milk and milk fat production were computed for the daughters in the two samples. Fourteen groups of sires, which were grouped according to number of daughters, were formed. While there was no restriction on the maximum number of sires that could fall in a group, a minimum of 30 sires in each group was allowed. The fourteenth group had only 28 sires.

All sires with less than 15 daughters or daughters in one herd only were excluded. Simple regressions of average of daughters in S_M on average of daughters in S_N were computed for all the fourteen groups and are listed, along with the values of other variables, in Table 16. The magnitude of regression, as expected, increases in general with increasing number of daughters in sample S_M of the group. In the first group, where the average number of daughters in the sample S_M is less than two, the regressions are very small, rather negative for both deviation milk and milk fat. Whereas, the highest regression for deviation milk was found to be 0.88, that of deviation fat was as high as 0.92.

Table 16. Values of variables involved in equation for regression of breeding value on daughter average

Group Number	Number of Sires	N^1	$X_{h1}=M$	$X_{h3}=4\frac{\Sigma m_i(m_i-1)}{M}$	Regressions; $Y_h=b_h$		k_h	
					Milk	Milk fat	Milk	Milk fat
1	162	21.25	1.96	3.447	-0.0291	-0.0260	37,089.4	13,939.1
2	62	16.87	6.35	8.914	0.2459	0.3045	12,537.6	3,977.3
3	47	11.96	9.53	13.339	0.5524	0.3821	6,143.4	2,431.9
4	41	10.46	11.53	18.486	0.4670	0.2902	5,381.7	1,998.8
5	56	10.80	14.11	27.716	0.7191	0.9149	6,544.2	2,524.2
6	51	11.86	17.06	35.770	0.6101	0.5798	7,769.3	2,821.9
7	55	15.07	20.16	33.125	0.6031	0.5383	7,079.2	2,519.7
8	56	17.43	25.07	43.744	0.6684	0.5173	6,079.8	2,164.5
9	43	17.65	31.79	62.841	0.6450	0.4640	6,475.1	2,335.9
10	37	29.05	39.92	50.210	0.8689	0.6168	4,639.0	1,804.9
11	33	50.09	56.39	29.876	0.4257	0.3943	4,482.5	1,573.3
12	32	80.28	79.91	25.405	0.8770	0.5572	3,188.2	1,176.4
13	30	117.47	117.47	20.629	0.7417	0.6174	3,285.4	1,256.1
14	28	251.68	251.07	23.271	0.8632	0.7086	3,495.2	1,077.3

[1]Number of daughters in sample S_N.

The data from Table 16 were first fitted to the non-linear equation 26 and then again after adjusting it for 28. The results of the convergence of least squares estimates in various cycles of iteration are given in Figures 4, 5, 6 and 7. The estimates $\hat{h}^2$ and $\hat{c}^2$ thus obtained are summarized in the first half of Table 17.

Table 17. Least squares estimates of h^2 and c^2 obtained by using Hartley's modified Gauss-Newton method

Trait	Simultaneous estimation		$\hat{h}^2$
	$\hat{h}^2$	$\hat{c}^2$	(c^2=0)
Before adjustment:			
Milk	0.257	-0.011	0.279
Milk fat	0.210	0.026	0.172
After adjustment:			
Milk	0.130	-0.048	0.178
Milk fat	0.126	-0.031	0.153

The values of $\hat{h}^2$ and $\hat{c}^2$ were reached within ten iterations. At least two different starting values were used and virtually the same final answers were reached in each computation. When the starting value was closed to the actual, relatively smaller cycles of iteration were needed to reach the final value.

The estimates of c^2, thus obtained, are considerably low. They are negative but small and can be considered equal to zero, for all traits, both before and after adjustment. But then estimates computed in 1 and 2

of section B were also considerably lower, though none were as low as these here.

For estimation of c^2 by Hartley's modified Gauss-Newton method, a completely theoretical approach is made. It is questionable whether this method yields results which are of any practical significance. No direct way of putting any confidence limits on the estimates thus computed, is available. An indirect approach is to compare the estimates of heritability obtained by this method to those computed by widely accepted procedures. The correlation among paternal half-sibs method was used to estimate heritability in section A. The estimates for deviation milk and milk fat were 0.232 and 0.188, respectively. Compared to these, estimates for both traits, obtained by using Hartley's procedure are slightly larger before the adjustment was made but are considerably smaller after the adjustment; as if without any adjustment, heritability is being over-estimated and after the adjustment it is underestimated.

It must be pointed out that while sires with less than 15 daughters were excluded from the analysis by iteration procedure, they were included in estimates of heritability from correlation among half-sib, as long as they had daughters in at least two herds. In the same section, it was found (Table 3) that when the sires with 2 or 3 daughters were dropped from the analyses, the heritability estimates also dropped down to 0.211 and 0.163, for deviation milk and milk fat, respectively.

Maybe the lower estimates of heritability, obtained after adjustment, can be explained on the basis that, because the sires with less daughters were excluded from the data, the estimates went down. Thus, the method does seem to have some significance and the estimates $\hat{h}^2$ and $\hat{c}^2$ are not

entirely unreliable.

Since estimates of c^2 were found to be so near zero, c^2 was assumed to be equal to zero and the data from Table 16 were fitted to the following model:

$$Y_h = f(x_h; \theta) = \frac{x_{h1}\,\theta_1}{4 + x_{h2}\,\theta_1} \qquad 29$$

which differs from 26 in only θ_2 i.e., c^2 term. Similar results of various cycles of iteration for heritability alone are present in Figures 8, 9, 10 and 11. The values of $\hat{h}^2$ so computed (second half of Table 17) are in agreement with those estimated from correlation among half-sibs, (0.178 and 0.153 vs. 0.211 and 0.163). This, indirectly, leads to a conclusion that the zero value of c^2, that is the extra correlation among paternal half-sibs in the same herd is probably valid at least in the present population.

An interesting feature about the authenticity of the procedure can be seen by comparing the two halves of Table 17. For milk fat, before adjustment, the estimate of c^2 was positive; when c^2 was dropped out of model, the estimate of h^2 went down. Whereas, in the other three cases the c^2 was negative, the estimate of h^2 went up for $c^2 = 0$.

In conclusion, these different methods do not estimate the same parameter. The use of intra-class correlations or heritability estimates among a sire's daughters in the same and different herds is a simple and straightforward approach to determine the magnitude of c^2. This c^2 is the difference of the intra-class correlations between paternal half-sibs in the same and different herds, and may be used to adjust records of those daughters that are in one herd. The values obtained by this method

DATA SET NO. 1
HARTLEY,S MODIFIED GAUSS-NEWTON METHOD, NOT ADJUSTED, CC & HH, MILK

ESTIMATE NO. 1 OF THE PARAMETERS

CYCLE	PARAMETER ESTIMATES AND ERROR SUM OF SQUARES		
1	0.500000000000D-01	0.150000000000D 00	0.841830981739D 00
2	-0.331952808478D-01	0.186792105022D 00	0.306806749520D 00
3	-0.133241094108D-01	0.246867174477D 00	0.289326882848D 00
4	-0.102998180258D-01	0.257593439150D 00	0.289007954505D 00
5	-0.106693864760D-01	0.257249364798D 00	0.289004259204D 00
6	-0.106300388462D-01	0.257378427869D 00	0.289004217890D 00
7	-0.106340853277D-01	0.257374776083D 00	0.289004217434D 00

SOLUTION VECTOR IS THETA(I):

-0.106336510-01
0.257376200 00

ESTIMATE NO. 2 OF THE PARAMETERS

CYCLE	PARAMETER ESTIMATES AND ERROR SUM OF SQUARES		
1	0.100000000000D-01	0.200000000000D 00	0.3940663349990 00
2	-0.131852786791D-01	0.241908955885D 00	0.290043760045D 00
3	-0.101283133762D-01	0.257454803944D 00	0.289016504440D 00
4	-0.106973886821D-01	0.257165044878D 00	0.289004328571D 00
5	-0.106286400355D-01	0.257377841340D 00	0.289004218518D 00
6	-0.106343164041D-01	0.257374105428D 00	0.289004217440D 00

SOLUTION VECTOR IS THETA(I):

-0.10633642D-01
0.25737619D 00

Figure 4. Convergence of least squares estimates of c^2 and h^2 for deviation milk in various cycles of iteration (before adjustment)

(The vector theta (I) has 2 elements CC = c^2 and HH = h^2; multiply each number by 10^k where k = last three digits of the number. The same goes for the following figures)

DATA SET NO. 2
HARTLEY,S MODIFIED GAUSS-NEWTON METHOD, NOT ADJUSTED, CC & HH, MILK FAT

ESTIMATE NO. 1 OF THE PARAMETERS

CYCLE	PARAMETER ESTIMATES AND ERROR SUM OF SQUARES		
1	0.500000000000D-01	0.1500000000000D 00	0.707900380782D 00
2	-0.216639472273D-02	0.1636664160660 00	0.596948074847D 00
3	0.235064023512D-01	0.2034271982190 00	0.5907017925120 00
4	0.262924164242D-01	0.2094538134820 00	0.5905731999390 00
5	0.261063554741D-01	0.2097662835380 00	0.590567607548D 00
6	0.261854537081D-01	0.2099812724000 00	0.5905673787900 00
7	0.261824640386D-01	0.2100001675830 00	0.590567369535D 00

SOLUTION VECTOR IS THETA(I):

0.26184945D-01
0.21000800D 00

ESTIMATE NO. 2 OF THE PARAMETERS

CYCLE	PARAMETER ESTIMATES AND ERROR SUM OF SQUARES		
1	0.300000000000D-01	0.2000000000000D 00	0.593473667246D 00
2	0.239911289320D-01	0.2042012412090 00	0.590682717782D 00
3	0.262437913099D-01	0.2094317633420 00	0.590572534767D 00
4	0.261179620496D-01	0.2097895079260 00	0.590567577913D 00
5	0.261842917628D-01	0.2099814250530 00	0.590567377626D 00

SOLUTION VECTOR IS THETA(I):

0.26182799D-01
0.21000094D 00

Figure 5. Convergence of least squares estimates of c^2 and h^2 for deviation milk fat in various cycles of iteration (before adjustment)

DATA SET NO. 3
HARTLEY,S MODIFIED GAUSS-NEWTON METHOD, ADJUSTED, CC & HH, MILK

ESTIMATE NO. 1 OF THE PARAMETERS

CYCLE	PARAMETER ESTIMATES AND ERROR SUM OF SQUARES		
1	0.5000000000000D-01	0.100000000000D 00	0.5892687356350 08
2	-0.1234424439760-02	0.913028879733D-01	0.454030572546D 08
3	-0.280859092554D-01	0.942066107864D-01	0.337608914383D 08
4	-0.504074198147D-01	0.125974079349D 00	0.2454779715380 08
5	-0.482133704722D-01	0.130615752208D 00	0.245023106733D 08
6	-0.481795080427D-01	0.130224562173D 00	0.245019572359D 08
7	-0.481473552888D-01	0.130280674436D 00	0.245019489867D 08

SOLUTION VECTOR IS THETA(I):

-0.48146565D-01
0.13027158D 00

ESTIMATE NO. 2 OF THE PARAMETERS

CYCLE	PARAMETER ESTIMATES AND ERROR SUM OF SQUARES		
1	0.3000000000000D-01	0.120000000000D 00	0.460511545961D 08
2	-0.121001924363D-01	0.108013989774D 00	0.356942844849D 08
3	-0.381277183411D-01	0.107462820279D 00	0.275080642786D 08
4	-0.473674582725D-01	0.130178472281D 00	0.2450799151200 08
5	-0.482158376009D-01	0.129797749035D 00	0.245022807617D 08
6	-0.481038942329D-01	0.130267479534D 00	0.245019665648D 08
7	-0.481495165365D-01	0.130247147477D 00	0.245019497507D 08
8	-0.481435194375D-01	0.130272420322D 00	0.245019488424D 08

SOLUTION VECTOR IS THETA(I):

-0.48145980D-01
0.13027132D 00

Figure 6. Convergence of least squares estimates of c^2 and h^2 for deviation milk in various cycles of iteration (after adjustment)

DATA SET NO. 4
HARTLEY,S MODIFIED GAUSS-NEWTON METHOD, ADJUSTED, CC & HH, MILK FAT

ESTIMATE NO. 1 OF THE PARAMETERS

CYCLE	PARAMETER ESTIMATES AND ERROR SUM OF SQUARES		
1	0.500000000000D-01	0.100000000000D 00	0.7013743532090 07
2	-0.5559218543880 02	0.9016020102190-01	0.5692986631510 07
3	-0.3894460241870-01	0.1119268959290 00	0.4630017301390 07
4	-0.3060113152780-01	0.1264613470550 00	0.4586450265590 07
5	-0.3097628525610-01	0.1259076357540 00	0.4586379210710 07
6	-0.3098359663480-01	0.1259214268370 00	0.4586379063950 07

SOLUTION VECTOR IS THETA(I):

-0.30983505D-01
0.125922780 00

ESTIMATE NO. 2 OF THE PARAMETERS

CYCLE	PARAMETER ESTIMATES AND ERROR SUM OF SQUARES		
1	0.100000000000D-01	0.120000000000D 00	0.5326890294860 07
2	-0.1591889572540-01	0.1112405022880 00	0.4870747371650 07
3	-0.3513580435660-01	0.1168870485830 00	0.4601091678250 07
4	-0.3043176501190-01	0.1257193740670 00	0.4586649673110 07
5	-0.3109107961870-01	0.1255558044160 00	0.4586401159180 07
6	-0.3093911598350-01	0.1259062687390 00	0.4586380840130 07
7	-0.3099234758860-01	0.1258929387140 00	0.4586379207850 07
8	-0.3098010694190-01	0.1259212673140 00	0.4586379075240 07
9	-0.3098442700100-01	0.1259201848790 00	0.4586379064480 07

SOLUTION VECTOR IS THETA(I):

-0.30983434D-01
0.125922480 00

Figure 7. Convergence of least squares estimates of c^2 and h^2 for deviation milk fat in various cycles of iteration (after adjustment)

DATA SET NO. 7

"HARTLEY'S MODIFIED GAUSS-NEWTON METHOD, NOT ADJUSTED, HH ONLY, MILK"

ESTIMATE NO. 1 OF THE PARAMETERS

CYCLE	PARAMETER ESTIMATES AND ERROR SUM OF SQUARES
1	0.150000000000D 00 0.517528108774D 00
2	0.236275474710D 00 0.3073734449780 00
3	0.2737053487110 00 0.290726976960 00
4	0.279296000214D 00 0.290452699818D 00
5	0.279638088595D 00 0.290451716020D 00

SOLUTION VECTOR IS THETA(I):

0.27965499D 00

ESTIMATE NO. 2 OF THE PARAMETERS

CYCLE	PARAMETER ESTIMATES AND ERROR SUM OF SQUARES
1	0.200000000000D 00 0.357188495904D 00
2	0.262449127147D 00 0.292852722020D 00
3	0.278225926537D 00 0.290467346603D 00
4	0.279582233680D 00 0.290451754852D 00
5	0.279652260013D 00 0.290451713717D 00

SOLUTION VECTOR IS THETA(I):

0.27965568D 00

Figure 8. Convergence of least squares estimate of h^2 for deviation milk in various cycles of iteration (before adjustment)

DATA SET NO. 8

"HARTLEY'S MODIFIED GAUSS-NEWTON METHOD, NOT ADJUSTED, HH ONLY, MILK FAT"

ESTIMATE NO. 1 OF THE PARAMETERS

CYCLE	PARAMETER ESTIMATES AND ERROR SUM OF SQUARES
1	0.150000000000D 00 0.6045534501900D 00
2	0.165915344780D 00 0.5960977296080D 00
3	0.170629146776D 00 0.5954799894710 00
4	0.171773393565D 00 0.5954454191240 00
5	0.172034241170D 00 0.5954436446400 00
6	0.172092786934D 00 0.5954435555010 00

SOLUTION VECTOR IS THETA(I):

0.17210588D 00

ESTIMATE NO. 2 OF THE PARAMETERS

CYCLE	PARAMETER ESTIMATES AND ERROR SUM OF SQUARES
1	0.200000000000D 00 0.6066002730950 00
2	0.176452737248D 00 0.5957476883680 00
3	0.173033032100D 00 0.5954575465090 00
4	0.172313816521D 00 0.5954442376440 00
5	0.172155158238D 00 0.5954435849660 00
6	0.172119810636D 00 0.5954435525140 00

SOLUTION VECTOR IS THETA(I):

0.172111920 00

Figure 9. Convergence of least squares estimate of h^2 for deviation milk fat in various cycles of iteration (before adjustment)

DATA SET NO. 5

"HARTLEY'S MODIFIED GAUSS-NEWTON METHOD, ADJUSTED, HH ONLY, MILK"

ESTIMATE NO. 1 OF THE PARAMETERS

CYCLE	PARAMETER ESTIMATES AND ERROR SUM OF SQUARES	
1	0.150000000000D 00	0.3211871040530 08
2	0.178090049379D 00	0.308840847100D 08
3	0.178381085203D 00	0.308839671486D 08

SOLUTION VECTOR IS THETA(I):

0.17838109D 00

ESTIMATE NO. 2 OF THE PARAMETERS

CYCLE	PARAMETER ESTIMATES AND ERROR SUM OF SQUARES	
1	0.200000000000D 00	0.314885784449D 08
2	0.178412283020D 00	0.308839684972D 08
3	0.178381093873D 00	0.308839671486D 08

SOLUTION VECTOR IS THETA(I):

0.17838109D 00

Figure 10. Convergence of least squares estimate of h^2 for deviation milk in various cycles of iteration (after adjustment)

DATA SET NO. 6

"HARTLEY'S MODIFIED GAUSS-NEWTON METHOD, ADJUSTED, HH ONLY, MILK FAT"

ESTIMATE NO. 1 OF THE PARAMETERS

CYCLE	PARAMETER ESTIMATES AND ERROR SUM OF SQUARES	
1	0.150000000000D 00	0.483846645453D 07
2	0.152943102458D 00	0.483670607140D 07

SOLUTION VECTOR IS THETA(I):

0.15294799D 00

ESTIMATE NO. 2 OF THE PARAMETERS

CYCLE	PARAMETER ESTIMATES AND ERROR SUM OF SQUARES	
1	0.200000000000D 00	0.521490573522D 07
2	0.152785282269D 00	0.483671137341D 07
3	0.152947974609D 00	0.483670606662D 07

SOLUTION VECTOR IS THETA(I):

0.15294799D 00

Figure 11. Convergence of least squares estimate of h^2 for deviation milk fat in various cycles of iteration (after adjustment)

indicate that c^2 of the different traits is of little importance because of their small magnitude. These values can be criticized because a large number of records are excluded from the analyses. This is particularly true in the analysis for daughters in the same herd since one herd per sire is selected at random while all his daughters that are in other herds, are dropped. If the daughters are in only a few herds, only one from each herd is used in the analysis for daughters in different herds.

Estimates of s, t and w demonstrate the environmental correlations due to herd and year-season effects separately as well as combined. These results agree very well with those by Thomson (1968) who used an independent set of data to compute these correlations. The values of s, t and w may be used to adjust the records of daughters in the same herd-year-season, herd and year-season, respectively. The additional correlation ρ^2 is a measure of common environments that paternal half-sibs in the same herd experience over and above the correlation due to common sire and common herd effects. To estimate ρ^2, the intra-class correlation among daughters in the same herd is used. This has a shortcoming of losing a great many records, especially when daughters are scattered in several herds. This is part of the reason to favor the estimates obtained by using daughters of single herd sires.

Intra-class correlation within herds and sires, computed from the deviation production of daughters of each sire, appears to estimate the environmental correlation which is due to the variation within herds that affects the paternal half-sibs in particular. The daughters of a certain sire may be treated differently in different herds. Therefore, the magnitude of c^2 may vary from herd to herd and from sire to sire. Even

so, for all practical purposes, a value of c^2 that can be used rather extensively, is needed. This procedure evidently yields results that are usable in the current USDA sire evaluation program. It was concluded that the magnitude of c^2 is large enough for it to be considered in estimating breeding values of sires. Also, that considering c^2 for AI sired daughters is as important as for NS sired daughters. The estimates obtained here are in agreement with the ones in current use.

Intuitively it is appealing to estimate c^2 and h^2 simultaneously because when c^2 alone is estimated, there is a possibility of a part of h^2 being included in the estimate. The same is true for estimation of h^2 alone. From this point of view, Hartley's iterative procedure seems important. The validity of the c^2 estimates obtained using this method is questionable, since they are negative. However, if the true value of c^2 is near zero, sampling variation could have led to negative estimates. The values obtained for h^2, on the other hand agree with those computed from the widely used method of intra-class correlation among half-sibs. The results from Hartley's method make one wonder if the value of c^2 is zero after all.

VI. SUMMARY

This study was undertaken to explore some of the possible approaches that can be used to estimate environmental correlation (c^2) among paternal half-sibs. Different procedures lead to different estimates of c^2 that are useful for different applications. It was expected to pinpoint some of the procedures that can be of practical importance. The data were obtained from Iowa Dairy Record Processing Center at Iowa State University and consisted of 55,170 first lactation records made in 2,326 herds located in eight midwestern states.

Estimates of heritability were obtained as background for judging later the c^2 and h^2 estimates. Heritability estimates and their standard errors for the four traits mature-equivalent (ME) milk and ME milk fat deviation (DEV) milk and DEV milk fat were 0.257 $\pm$ 0.019, 0.200 $\pm$ 0.017, 0.211 $\pm$ 0.018 and 0.163 $\pm$ 0.016, respectively. When sires with less than 4 daughters in at least two herds were included in the analyses, the corresponding estimates increased to, 0.284 $\pm$ 0.020, 0.240 $\pm$ 0.020, 0.232 $\pm$ 0.020 and 0.188 $\pm$ 0.016, respectively.

Four methods were used to compute c^2 estimates. The method of difference in intra-class correlations among daughters in one herd versus in many herds is based on the fact that the value of c^2 is maximum in the former case but is zero when each daughter is in a different herd. The estimates obtained were 0.014, 0.046 and 0.016 and 0.036 for the four traits, respectively. One statistical shortcoming of the method is that only a small portion of data is used.

The method of intra-class correlations due to herd-year-seasons, herds and year-seasons leads to three estimates of environmental

correlations, s, t and w, respectively. The estimates of environmental correlations (s) among daughters in the same herd and year-season for the four traits, were 0.357 ± 0.005, 0.381 ± 0.005, 0.045 ± 0.004 and 0.060 ± 0.004, respectively. For records made in the same herd but different year-seasons, the environmental correlations (t) were 0.199 ± 0.009, 0.216 ± 0.008, 0.038 ± 0.006 and 0.043 ± 0.006, respectively. The environmental correlations (w), among daughters in same year-season but different herds were 0.022 ± 0.007, 0.018 ± 0.006, 0.015 ± 0.005 and 0.010 ± 0.003, for the four traits, respectively. All the correlations for deviation records are above zero indicating that deviating records did not remove herd- and year-season effects completely. The additional environmental correlation (ρ^2) was estimated for each type of record. This correlation is a measure of common environments that paternal half-sibs in the same herd experience, over and above the correlation due to common sire and common herd effects. These estimates were 0.064, 0.095, -0.010 and 0.002, respectively. A large number of records are excluded from the analyses to get the intra-class correlation among half-sibs in the same herd. To use a greater proportion of the data, sires whose daughters were in a single herd were used. The corresponding estimates were 0.059, 0.068, 0.016 and 0.024, respectively.

The method of intra-class correlations within herds and sires computed from the deviation records leads to estimation of environmental correlation which is due to the within herd variation that affects those paternal half-sibs in particular. These correlations were computed for DEV milk and DEV milk fat and were 0.103 ± 0.007 and 0.113 ± 0.007, respectively among AI sired daughters; and 0.112 ± 0.015 and 0.120 + 0.015,

respectively for NS sired daughters. These were computed in such a way as to be directly applicable to the current USDA sire evaluation program. They are approximately of the same magnitude as the 0.140 value currently in use. It was concluded that the magnitude of c^2 is large enough for it to be considered in computation of breeding values. Also, that considering c^2 for AI sired daughters is as important as for NS sired daughters.

For the method of simultaneous estimation of c^2 and h^2, Hartley's modified Gauss-Newton method was used. A formula for the regression of the average of a sample of daughters on the average of an independent sample of daughters of the same sire, was developed which included the terms c^2, h^2 and a term to account for the distribution of daughters across herds. This was rewritten in the form of a non-linear model to which Hartley's iteration method was applied. The estimates of c^2 for DEV milk fat were -0.011 and 0.026, respectively; and the respective h^2 estimates were 0.257 and 0.210. An adjustment was made to correct for the random error in estimation of dependent variable in the regression model. The corresponding values of c^2 for the two traits after adjustment, were -0.048 and -0.031, respectively; the respective h^2 estimates were 0.130 and 0.126. The validity of these c^2 estimates is questionable from a practical point of view since they are negative. However, if the true value of c^2 is near zero, sampling variation could have led to negative estimates. Assuming $c^2 = 0$, estimates for h^2 alone were computed. For the two traits they were 0.273 and 0.172, respectively before adjustment and 0.178 and 0.153, respectively after adjustment. These estimates of h^2 do agree with those computed from the method of intra-class correlation among paternal half-sibs.

VII. LITERATURE CITED

Allaire, F. R. and S. N. Gaunt. 1965. First lactation contemporary comparisons as indicators of environmental influences on daughter records used for sire evaluation. J. Dairy Sci. 48: 454-461.

Atkinson, Joe D. 1966. TARSIER (FitTing non-lineAr regRession functions by leaSt squares, using the modIfied Gauss-Newton mEthod by H. O. HaRtley): Reference manual. Iowa State University, Statistical Laboratory, Numerical Analysis Programming Series No. 8.

Bereskin, B. 1963. Effects of genetic and environmental variance on diary sire evaluation. Unpublished Ph.D. thesis. Ames, Iowa, Library, Iowa State University of Science and Technology.

Bereskin, B. and A. E. Freeman. 1965. Genetic and environmental factors in dairy sire evaluation. I. Effects of herds, months and year-seasons on variance among lactation records: repeatability and heritability. J. Dairy Sci. 48: 347-359.

Bereskin, B. and J. L. Lush. 1965. Genetic and environmental factors in dairy sire evaluation. III. Influence of environmental and other extraneous correlations among the daughters. J. Dairy Sci. 48: 356-360.

Blanchard, R. P. 1965. Variation in the lactation yield of milk constituents. Unpublished M.S. thesis. Ames, Iowa, Library, Iowa State University of Science and Technology.

Blanchard, R. P., A. E. Freeman and P. W. Spike. 1966. Variation in lactation yield of milk constituents. J. Dairy Sci. 49: 953-956.

Burdick, J. M. and L. D. McGilliard. 1963. Interactions between sires in artificial insemination and management of dairy herds. J. Dairy Sci. 46: 452-458.

Butcher, D. F. 1965. The effect of selection on estimates of genetic parameters. Unpublished Ph.D. thesis. Ames, Iowa, Library, Iowa State University of Science and Technology.

Butcher, D. F. and A. E. Freeman. 1968. Heritabilities and repeatabilities of milk and milk fat production by lactation. J. Dairy Sci. 51: 1387-1391.

Gaunt, S. N. and J. E. Legates. 1958. Relative merits of five measures of dairy sire's transmitting ability. J. Dairy Sci. 830-839.

Hartley, H. O. 1961. Modified Gauss-Newton method for the fitting of non-linear regression functions by least squares. Technometrics 3: 269-280.

Heidhues, T., L. D. VanVleck, and C. R. Henderson. 1961. Actual and expected accuracy of sire proofs under the New York system of sampling bulls. Zeitschrift für Tierzüchtung and Zuchtüngsbiologis 75: 323-330.

Hickman, C. G. and C. R. Henderson. 1955. Components of the relationship between level of production and rate of maturity in dairy cattle. J. Dairy Sci. 38: 883-890.

Johansson, I. 1960. Progeny testing methods in Europe. J. Dairy Sci. 43: 706-713.

Johansson, I. 1961. Genetic aspects of dairy cattle breeding. Urbana, Illinois, University of Illinois Press.

Kelleher, D. J. 1964. The importance of bull x herd-year-season interaction in milk production. Unpublished Ph.D. thesis. Ames, Iowa, Library, Iowa State University of Science and Technology.

Kelleher, D. J., A. E. Freeman and J. L. Lush. 1967. Importance of bull x herd-year-season interaction in milk production. J. Dairy Sci. 50: 1703-1707.

Kendrick, J. F. 1955. Standardizing D.H.I.A. records in proving sires. United States Department of Agriculture (Publication) ARS-52-1.

Legates, J. E., F. J. Verlinden, and J. F. Kendrick. 1956. Sire by herd interaction in production traits in dairy cattle. J. Dairy Sci. 39: 1055-1063.

Lush, J. L. 1931. The number of daughters necessary to prove a sire. J. Dairy Sci. 14: 209-220.

Lush, J. L. 1935. Progeny test and individual performance as indicators of an animal's breeding value. J. Dairy Sci. 18: 1-19.

Lush, J. L. and L. D. McGilliard. 1955. Proving dairy sires and dams. J. Dairy Sci. 38: 163-180.

Mason, I. L. and A. Robertson. 1956. Progeny testing of dairy bulls at different levels of productions. J. Agric. Sci. 47: 367-375.

McDaniel, B. T. and E. L. Corley. 1965. Reliability of early sire evaluations. (Abstract) J. Dairy Sci. 48: 803.

Meek, A. M. and L. D. Van Vleck. 1964. Relationships between sire proofs. J. Dairy Sci. 47: 642-645.

Miller, R. H. and E. L. Corley. 1965. Usefulness of information on mates of sires in artificial insemination. J. Dairy Sci. 48: 580-585.

Plowman, R. D. and B. T. McDaniel. 1968. Changes in U.S.D.A. sire summary procedures. J. Dairy Sci. 51: 306-311.

Ramsay, J. M. 1966. Heterogeneity of variance at different genetic levels of milk production. Unpublished Ph.D. thesis. Ames, Iowa, Library, Iowa State University of Science and Technology.

Robertson, A., L. K. O'Connor, and J. Edwards. 1960. Progeny testing dairy bulls at different management levels. Anim. Prod. 2: 141-152.

Swiger, L. A., W. R. Harvey, D. O. Everson, and K. E. Gregory. 1964. The variance of intra-class correlation involving groups with one observation. Biometrics 20: 818-826.

Thomson, G. M. 1968. Selection of sires for use in artificial breeding. Unpublished Ph.D. thesis. Ames, Iowa, Library, Iowa State University of Science and Technology.

Touchberry, R. W., K. Rottensten and H. Andersen. 1960. A comparison of dairy sire progeny tests made at special Danish testing stations with tests made in farmer herds. J. Dairy Sci. 43: 529-545.

Tucker, W. L. and J. E. Legates. 1965. Seasonal division of herd-mates in sire evaluation. J. Dairy Sci. 48: 234-242.

Van Vleck, L. D. 1963. Genotype and environments in sire evaluation. J. Dairy Sci. 46: 983-987.

Van Vleck, L. D. and G. E. Brandford. 1965. Comparison of heritability estimates from daughter-dam regression and paternal half-sib correlation. J. Dairy Sci. 48: 1372-1375.

Van Vleck, L. D. 1966a. Heritability estimates of milk production with different number of records per sire by herd sub-class. J. Dairy Sci. 49: 53-55.

Van Vleck, L. D. 1966b. Environmental correlations in sire evaluation. J. Dairy Sci. 49: 56-60.

Van Vleck, L. D. 1966c. Paternal half-sib correlations between pairs in the same and different herds. J. Dairy Sci. 195-198.

Van Vleck, L. D. and C. L. Hart. 1965. Heritability estimates when dams and daughters are in the same and different herds. J. Dairy Sci. 48: 1676-1679.

Van Vleck, L. D., L. H. Wadell and C. R. Henderson. 1961. Components of variance associated with milk and milk fat records of artificially sired Holstein daughters. J. Anim. Sci. 20: 812-816.

Wadell, L. H. and L. D. McGilliard. 1959. Influence of artificial breeding on production of Michigan dairy herds. J. Dairy Sci. 42: 1079-1085.

Wright, S. 1921. Systems of mating. I. The biometric relations between parents and offspring. Genetics 16: 111-123.

Wright, S. 1934. The methods of path coefficients. Annals of Mathematical Statistics 5: 161-215.

VIII. ACKNOWLEDGMENTS

I wish to express my sincere appreciation to Dr. A. E. Freeman for his advice and counsel throughout the course of my graduate study. Without his gracious encouragement and outstanding guidance this manuscript could not have been possible. Also, I am grateful for his imparting to me the practical aspects of dairy science research. Only he could create a 'farmboy' out of a 'city slicker'.

Thanks are also due to Drs. L. N. Hazel, D. F. Cox, W. F. Hollander and Edward Pollak for their comments and criticism in reviewing the manuscript and serving on my graduate committee. Special thanks to Dr. George Zyskind for contributing several stimulating ideas and finally taking a place on my graduate committee.

I would also like to express my appreciation to Dr. Basil Eastwood for his help in explaining the data and to the fellow graduate students for their time to time discussions, comments and suggestions.

Last, but not least, to my typist Mrs. Carol Hill, "thank you for enduring me and my darn thesis".